

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 16:41:19 ; Search time 60 seconds
(without alignments)
1582.266 Million cell updates/sec

Title: US-10-650-369-12

Perfect score: 1715
Sequence: 1 MVKVGINGRIGRLAFRR.....EMSYAQLVRLLEYFAKIAK 336

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1715	100.0	336	5	AAM50665 Streptoco
2	1715	100.0	336	5	AAM50639 Streptoco
3	1714	99.9	336	5	ABP29960 Streptoco
4	1714	99.9	336	6	ABU46455 Protein e
5	1714	99.9	345	5	ABP29106 Streptoco
6	1711	99.8	336	2	AAR56486 Plasmin r
7	1711	99.8	336	4	AAV85681 Streptoco
8	1656.5	96.6	448	5	AAM50664 Streptoco
9	1618	94.3	336	5	AAM50667 Streptoco
10	1618	94.3	336	5	AAM50641 Streptoco
11	1609	93.8	336	5	ABP30758 Streptoco
12	1602	93.4	336	5	AAM50666 Streptoco
13	1602	93.4	336	5	AAM50640 Streptoco
14	1593.5	92.9	335	6	ABU02516 S. pneumo
15	1593.5	92.9	359	4	AAU37576 Streptoco
16	1593.5	92.9	359	6	ABU46262 Protein e
17	1592.5	92.9	359	4	AAU38000 Streptoco
18	1584	92.4	336	5	AAM50669 Streptoco
19	1584	92.4	336	5	AAM50643 Streptoco
20	1579.5	92.1	337	6	ABU44330 Protein e
21	1578	92.0	336	5	AAM50668 Streptoco
22	1578	92.0	336	5	AAM50642 Streptoco
23	1549.5	90.3	333	2	AAW55089 Streptoco
24	1549.5	90.3	333	2	ABP54583 S. pneumo
25	1549.5	90.3	333	7	ADC45135 S. pneumo

26	1410	82.2	336	5	ABB55601
27	1337.5	78.0	333	4	AAU35254
28	1337.5	78.0	333	6	ABU29370 Protein e
29	1333.5	77.8	336	5	ABB53868
30	1325.5	77.3	333	7	ADC95142
31	1277	74.5	335	6	ABU25342
32	1262	73.6	336	5	ABB48810
33	1262	73.6	336	6	ABU32568
34	1259.5	73.4	334	6	ABU23595
35	1217.5	71.0	335	6	ABU43810
36	1197.5	69.8	334	6	ABU37737
37	1196.5	69.8	357	6	ABP81005
38	1196.5	69.8	357	6	ABU37077
39	1190	69.4	336	4	AAG83058
40	1190	69.4	336	6	ABU42966
41	1190	69.4	348	5	ABP40668
42	1183	69.0	336	4	AAU36947
43	1183	69.0	336	4	AAU34122
44	1183	69.0	336	6	ABJ19195
45	1183	69.0	336	6	ABU16073

ALIGNMENTS

RESULT 1

AAM50665
ID AAM50665 standard; protein; 336 AA.

XX	AC	AAM50665;
XX	DT	08-APR-2002 (first entry)
XX	DE	Streptococcus dysgalactiae gapC plasmin binding protein DysGapC.
XX	DY	DysGapC; CapC; plasmin binding protein; epitope; infection; vaccine; immunisation; mastitis; therapy.
XX	OS	Streptococcus dysgalactiae.
XX	PN	WO200196379-A2.
XX	PD	20-DEC-2001.
XX	PF	11-JUN-2001; 2001WO-CA000836.
XX	PR	12-JUN-2000; 2000US-0211247P.
XX	PA	(UYSA-) UNIV SASKATCHEWAN.
XX	PI	Potter AA, Perez-Casal J, Fontaine M;
XX	DR	WPI; 2002-098051/13.
XX	PT	N-PSDB; ABA91328.
XX	PS	Novel GapC multiple epitope fusion polypeptide comprising antigenic determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S. parauberis, or S. iniae GapC protein useful for treating mastitis in vertebrates.
XX	CC	Claim 8; Fig 1A-B; 116pp; English.

The present sequence is that of the GapC plasmin binding protein, DysGapC, of Streptococcus dysgalactiae ATCC 43078, an isolate from a case of bovine mastitis. The invention relates to novel GapC multiple epitope fusion proteins that comprise epitopes from 1 or more of Streptococcus dysgalactiae, Streptococcus agalactiae, Streptococcus parauberis, Streptococcus uberis and Streptococcus iniae (see AAM50665-69). A claimed example is PolyGap4 (see AAM50664). Expression vectors and host cells for production of the multiple epitope fusion proteins are provided. The multiple epitope proteins are used in claimed vaccines for treating or preventing a bacterial infection in a vertebrate, especially a streptococcal infection, and particularly mastitis. They are also used in

CC claimed methods of detecting Streptococcus antibodies. The multiple
 CC epitope proteins are capable of eliciting broad immunity against a
 CC variety of streptococcal infections while minimising the number of
 CC antigens present in the final formulation and concomitantly reducing
 CC production costs
 XX
 SQ Sequence 336 AA;
 Query Match 100.0%; Score 1715; DB 5; Length 336;
 Best Local Similarity 100.0%; Pred. No. 8.6e-145; Indels 0; Gaps 0;
 Matches 336; Conservative 0; Mismatches 0;
 QY 1 MVVKVINGFGRIQLAFRRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
 DB 1 MVVKVINGFGRIQLAFRRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
 QY 61 KEGGFEVNGNFIKVSARDPENIDWATGVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
 DB 61 KEGGFEVNGNFIKVSARDPENIDWATGVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
 QY 121 TAPGGNDVTVVFNTHDILDTETVIGASCTTNCPLAPMAKALHDAFGIQKGLMTTIHA 180
 DB 121 TAPGGNDVTVVFNTHDILDTETVIGASCTTNCPLAPMAKALHDAFGIQKGLMTTIHA 180
 QY 181 YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
 DB 181 YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
 QY 241 GSVTELVVTLTKNVSVDENIAAKAASNDSPGYTDEDPVSSDIVGVSGLFDATQTKM 300
 DB 241 GSVTELVVTLTKNVSVDENIAAKAASNDSPGYTDEDPVSSDIVGVSGLFDATQTKM 300
 QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
 DB 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
 RESULT 2
 AAM50639
 ID AAM50639 standard; protein; 336 AA.
 XX
 AC AAM50639;
 XX
 DT 04-APR-2002 (first entry)
 XX
 DE Streptococcus dysgalactiae gapC plasmin binding protein.
 XX
 KW GapC; plasmin-binding protein; DysgalGapC; infection; mastitis; vaccine;
 KW diagnosis; therapy.
 XX
 OS Streptococcus dysgalactiae.
 XX
 PN WO200196381-A2.
 XX
 PD 20-DEC-2001.
 XX
 PF 11-JUN-2001; 2001WO-CA000838.
 XX
 PR 12-JUN-2000; 2000US-0211022P.
 XX
 PA (UYSA-) UNIV SASKATCHEWAN.
 XX
 PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;
 PI WPI; 2002-130725/17.
 DR N-PSDB; ABA91248.
 DR
 PT Novel isolated GapC protein of Streptococcus dysgalactiae, S.agalactiae,
 PT S.uberis, S.parauheris, or S.iniae, useful as vaccine component for
 PT treating streptococcal infection which causes mastitis in vertebrates.
 XX
 PS Claim 1(a); Fig 1A-B; 107pp; English.
 XX

CC The present sequence is that of the GapC plasmin binding protein
 CC (DysgalGapC) of Streptococcus dysgalactiae ATCC 43078, a clinical isolate
 CC from bovine mastitis. It is encoded by the GapC gene given in ABA91248.
 CC GapC protein, which has no signal sequence or membrane anchor domain, is
 CC capable of eliciting an immune response in a vertebrate. The invention
 CC provides the GapC genes and proteins of 5 Streptococcus species, as well
 CC as recombinant vectors, host cells and vaccine compositions comprising
 CC GapC polynucleotides or proteins. The vaccines are used to treat or
 CC prevent a bacterial infection, especially a streptococcal infection, and
 CC mastitis in particular (claimed). GapC proteins are also used in claimed
 CC methods for detecting GapC antibodies, and to raise antibodies that are
 CC used in claimed methods for detecting GapC proteins
 XX
 SQ Sequence 336 AA;
 Query Match 100.0%; Score 1715; DB 5; Length 336;
 Best Local Similarity 100.0%; Pred. No. 8.6e-145; Indels 0; Gaps 0;
 Matches 336; Conservative 0; Mismatches 0;
 QY 1 MVVKVINGFGRIQLAFRRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
 DB 1 MVVKVINGFGRIQLAFRRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
 QY 61 KEGGFEVNGNFIKVSARDPENIDWATGVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
 DB 61 KEGGFEVNGNFIKVSARDPENIDWATGVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
 QY 121 TAPGGNDVTVVFNTHDILDTETVIGASCTTNCPLAPMAKALHDAFGIQKGLMTTIHA 180
 DB 121 TAPGGNDVTVVFNTHDILDTETVIGASCTTNCPLAPMAKALHDAFGIQKGLMTTIHA 180
 QY 181 YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
 DB 181 YTGDMILDGPHRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
 QY 241 GSVTELVVTLTKNVSVDENIAAKAASNDSPGYTDEDPVSSDIVGVSGLFDATQTKM 300
 DB 241 GSVTELVVTLTKNVSVDENIAAKAASNDSPGYTDEDPVSSDIVGVSGLFDATQTKM 300
 QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
 DB 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
 RESULT 3
 ABP29960
 ID ABP29960 standard; protein; 336 AA.
 XX
 AC ABP29960;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 9096.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX
 OS Streptococcus pyogenes.
 XX
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GB004789.
 XX
 PR 27-OCT-2000; 2000GB-00026333.
 PR 24-NOV-2000; 2000GB-00028727.
 PR 07-MAR-2001; 2001GB-00005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX

PI Telford J, Massignani V, Margarit Y Rosl, Grandi G, Fraser C;
 PI Tettelin H;
 XX WPI; 2002-352536/38.
 DR N-PSDB; ABN70591.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection or
 PT disease caused by Streptococcus bacteria, such as meningitis, and for
 PT detecting a compound that binds to the protein.
 PS
 XX Claim 1; Page 4031; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (II), Nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 XX
 SQ Sequence 336 AA;

Query Match 99.9%; Score 1714; DB 5; Length 336;
 Best Local Similarity 99.7%; Pred. No. 1.1e-144;
 Matches 335; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRFDGTV 60
 DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRFDGTV 60

QY 61 KEGGFVNGNFIKVSARDPENIDWATDGVIVLEATGFFAKKAAEKHLHANGAKKVI 120
 DB 61 KEGGFVNGNFIKVSARDPENIDWATDGVIVLEATGFFAKKAAEKHLHANGAKKVI 120

QY 121 TAPGNDVKTVVNTNHDILDTETVSGACTTNCNCLAPMAKALHDAFGKGLMTTIIA 180
 DB 121 TAPGNDVKTVVNTNHDILDTETVSGACTTNCNCLAPMAKALHDAFGKGLMTTIIA 180

QY 181 YTGDMILDGPHRGDLRRARAGAAIIVPNSGTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
 DB 181 YTGDMILDGPHRGDLRRARAGAAIIVPNSGTGAAGAIGLVIPELNGKLDGAAQRPVPT 240

QY 241 GSVTELVTLDKNVSVDEINAMKAASNDSPGYTDEPTVSSDIVGVSGLFDATQTKVM 300
 DB 241 GSVTELVTLDKNVSVDEINAMKAASNDSPGYTDEPTVSSDIVGVSGLFDATQTKVM 300

QY 301 EYDGSQLVKVSVDNENSYTAQLVRLTLEYFAKIAK 336
 DB 301 EYDGSQLVKVSVDNENSYTAQLVRLTLEYFAKIAK 336

RESULT 4
 ABU46455
 ID ABU46455 standard; protein; 336 AA.
 XX
 AC ABU46455;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #31982.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX

OS Streptococcus pyogenes.
 XX WO200277183-A2.
 PN
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 PA
 XX Wang L, Zamudio C, Malone C, Hasebeck R, Ohlsen KL, Zyskind JW;
 PI Walli D, Tzawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 PI
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA50325.
 DR
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 74379; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than S. aureus, S. typhimurium,
 CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 336 AA;

Query Match 99.9%; Score 1714; DB 6; Length 336;
 Best Local Similarity 99.7%; Pred. No. 1.1e-144;
 Matches 335; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRFDGTV 60
 DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRFDGTV 60

QY 61 KEGGFVNGNFIKVSARDPENIDWATDGVIVLEATGFFAKKAAEKHLHANGAKKVI 120
 DB 61 KEGGFVNGNFIKVSARDPENIDWATDGVIVLEATGFFAKKAAEKHLHANGAKKVI 120

CC used in gene therapy. Antibodies to (I) are used for affinity chromatography, immunoassays, and distinguishing/identifying Streptococcus proteins

CC Sequence 345 AA;

Query Match 99.9%; Score 1714; DB 5; Length 345;
Best Local Similarity 99.7%; Pred. No. 1.1e-144;
Matches 335; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVYKVGINGFGRIQRLAFRRIONVEGVETRLNDLTDNMLAHLKYDTTQGRFDGTVEV 60
Db 10 MVYKVGINGFGRIQRLAFRRIONVEGVETRLNDLTDNMLAHLKYDTTQGRFDGTVEV 69
QY 61 KEGGFVNGNFKVSAERDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 70 KEGGFVNGNFKVSAERDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVI 129
QY 121 TAPGGNDVTVVNTNHDILDTETVISCSTTNCLAPMAKALHDAFGIQGLMTTIIHA 180
Db 130 TAPGGNDVTVVNTNHDILDTETVISCSTTNCLAPMAKALHDAFGIQGLMTTIIHA 189
QY 181 YTGDMILDPHRRGDLRRARAGANIVPNSGAAKAGLVIPELNGKLDGAAQRVFPT 240
Db 190 YTGDMILDPHRRGDLRRARAGANIVPNSGAAKAGLVIPELNGKLDGAAQRVFPT 249
QY 241 GSVELVTLTKNSVDEINAAKASNDSPGYTEDPIVSSDIVGSYGLSFDATQTKM 300
Db 250 GSVELVTLTKNSVDEINAAKASNDSPGYTEDPIVSSDIVGSYGLSFDATQTKM 309
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 310 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 345

RESULT 6
AAR56486
ID AAR56486 standard; protein; 336 AA.
AC AAR56486;
XX
DT 25-MAR-2003 (revised)
DT 22-FEB-1995 (first entry)
XX
DE Plasmin receptor.
XX
KW Plasmin receptor; isolate 64/14; plasmin; alpha-2-antiplasmin;
KW plasminogen activator; bleeding; reocclusion; thrombosis;
KW pulmonary embolism; clots.
XX
OS Streptococcus pyogenes.
XX
PN US5328996-A.
XX
PD 12-JUL-1994.
XX
PF 10-AUG-1992; 92US-00928462.
XX
PR 29-MAR-1989; 89US-00330849.
PR 16-MAY-1990; 90US-00524411.
XX
XX (UYFL) UNIV FLORIDA RES FOUND INC.
XX
PI Von Mering G, Broder C, Boyle MDP, Lottenberg R;
XX
XX WPI; 1994-225327/27.
XX
DR N-PSDB; AAQ70705.
XX
PT New DNA encoding bacterial plasmin receptor - useful as thrombolytic agents, used with plasminogen activator or bound to plasmin, also useful in vaccines against bacterial infection.
XX
XX Claim 1; Col 27-30; 19pp; English.

QY 121 TAPGGNDVTVVNTNHDILDTETVISCSTTNCLAPMAKALHDAFGIQGLMTTIIHA 180
Db 121 TAPGGNDVTVVNTNHDILDTETVISCSTTNCLAPMAKALHDAFGIQGLMTTIIHA 180
QY 181 YTGDMILDPHRRGDLRRARAGANIVPNSGAAKAGLVIPELNGKLDGAAQRVFPT 240
Db 181 YTGDMILDPHRRGDLRRARAGANIVPNSGAAKAGLVIPELNGKLDGAAQRVFPT 240
QY 241 GSVELVTLTKNSVDEINAAKASNDSPGYTEDPIVSSDIVGSYGLSFDATQTKM 300
Db 241 GSVELVTLTKNSVDEINAAKASNDSPGYTEDPIVSSDIVGSYGLSFDATQTKM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 5
ABP29106
ID ABP29106 standard; protein; 345 AA.
XX
AC ABP29106;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 7388.
XX
KW Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus pyogenes.
XX
PN WO200234771-A2.
PD
PF 29-OCT-2001; 2001WO-GB004789.
XX
XX 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX
XX (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
XX Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C;
PI Tettelin H;
XX
XX WPI; 2002-352536/38.
DR N-PSDB; ABN69737.
XX
XX New Streptococcus protein for the treatment or prevention of infection or disease caused by Streptococcus bacteria, such as meningitis, and for detecting a compound that binds to the protein.
XX
XX Claim 1; Page 3888; 4525pp; English.

The invention relates to a protein (ABP25413-ABP30895) from group B streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in the specification. The proteins have antibacterial and antiinflammatory activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and antibodies that bind (I) are used in the manufacture of medicaments for the treatment or prevention of infection or disease caused by Streptococcus bacteria, particularly S. agalactiae and S. pyogenes. Nucleic acids encoding (I) are used to detect Streptococcus in a biological sample. (I) is used to determine whether a compound binds to (I). A composition comprising (I) or a nucleic acid encoding (I), may be used as a vaccine or diagnostic composition. The disease caused by Streptococcus that is prevented or treated may be meningitis. Nucleic acid encoding (I) may be used to recombinantly produce (I) and may be

XX This sequence represents the S. pyogenes plasmin receptor. The DNA
CC encoding this sequence was isolated from the S. pyogenes clinical isolate
CC 64/14. The plasmin receptor has a very high affinity for plasmin which,
CC when bound, retains its enzymatic activity but is not regulated
CC (inhibited) by alpha-2-antiplasmin. The receptor protein, when
CC administered concurrently or sequentially, prolongs the activity of
CC plasminogen activator (pA) so allows a reduction in dose, and thus lowers
CC the risk of bleeding, and may prevent reocclusion of blood vessels. The
CC protein may be coupled to a fibrin-specific monoclonal antibody to
CC provide targeting to clots. The plasmin receptor may be useful in human
CC or veterinary medicine, for treatment of thrombosis and pulmonary
CC embolism, and for solubilising clots in catheters or shunts. (Updated on
CC 25-MAR-2003 to correct PF field.)
XX SQ Sequence 336 AA;

Query Match 99.8%; Score 1711; DB 2; Length 336;
Best Local Similarity 99.4%; Pred. No. 2e-144; Mismatches 0; Indels 0; Gaps 0;
Matches 334; Conservative 2;

QY 1 MVVKVINGFGRIQLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
DB 1 MVVKVINGFGRIQLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120
DB 61 KEGGFVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTVPVNTNHDILDTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 180
DB 121 TAPGNDVKTVPVNTNHDILDTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 180
QY 181 YTGDMILDGHRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAQRVPVPT 240
DB 181 YTGDMILDGHRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAQRVPVPT 240
QY 241 GSVTELVTLDKNVSDVEINAAKASNDSPGYTEDPIVSSDIVGVSGLFDATQTKM 300
DB 241 GSVTELVTLDKNVSDVEINAAKASNDSPGYTEDPIVSSDIVGVSGLFDATQTKM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336
DB 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336

RESULT 7
AA85681
ID AA85681 standard; protein; 336 AA.
XX AC AA85681;
XX DE 13-FEB-2001 (first entry)
XX DE Streptococcal plasmin receptor amino acid sequence.
XX KW Plasmin receptor; plr; immune response; vaccination; antibacterial;
XX KW Streptococcal infection; antihelminthic.
XX OS Streptococcus pyogenes.
XX PN US6136323-A.
XX PD 24-OCT-2000.
XX PF 11-JUL-1994; 94US-00273247.
XX PR 29-MAR-1989; 89US-00330849.
XX PR 16-MAY-1990; 90US-00524411.
XX PR 10-AUG-1992; 92US-00928462.
XX PA (UYFL) UNIV FLORIDA RES FOUND INC.
XX

PI Von Mering G, Broder C, Lottenberg R, Boyle MDP;
XX WPI; 2001-006210/01.
DR N-PSDB; AAC66140.
XX Raising an immune response in a mammal, especially for vaccination
PT against group A streptococcal infections, comprises administering an
PT isolated streptococcal plasmin receptor protein.
XX Claim 2; Col 27-30; 17pp; English.
XX Polynucleotide sequence AAC66140 encodes the Streptococcal plasmin
CC receptor protein AAY85681. The protein is used in a method for raising an
CC immune response in a mammal. The method comprises administering the
CC plasmin receptor protein. The method is useful as a vaccination against
CC group A Streptococcal infections and potentially against a broad range of
CC infections associated with pathogens expressing glyceraldehyde-3-
CC phosphate dehydrogenase (GAPDH) or GAPDH-like proteins at their cell
CC surfaces. The vaccine has antibacterial and antihelminthic activity
XX SQ Sequence 336 AA;

Query Match 99.8%; Score 1711; DB 4; Length 336;
Best Local Similarity 99.4%; Pred. No. 2e-144; Mismatches 0; Indels 0; Gaps 0;
Matches 334; Conservative 2;

QY 1 MVVKVINGFGRIQLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
DB 1 MVVKVINGFGRIQLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120
DB 61 KEGGFVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTVPVNTNHDILDTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 180
DB 121 TAPGNDVKTVPVNTNHDILDTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 180
QY 181 YTGDMILDGHRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAQRVPVPT 240
DB 181 YTGDMILDGHRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAQRVPVPT 240
QY 241 GSVTELVTLDKNVSDVEINAAKASNDSPGYTEDPIVSSDIVGVSGLFDATQTKM 300
DB 241 GSVTELVTLDKNVSDVEINAAKASNDSPGYTEDPIVSSDIVGVSGLFDATQTKM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336
DB 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336

RESULT 8
AAM50664
ID AAM50664 standard; protein; 448 AA.
XX AC AAM50664;
XX AC AAM50664;
XX DT 29-AUG-2003 (revised)
XX DT 08-APR-2002 (first entry)
XX DE Streptococcus GapC multiepitope fusion PolyGap4.
XX KW PolyGap4; GapC; plasmin binding protein; epitope; infection; vaccine;
XX KW immunisation; mastitis; therapy.
XX OS Streptococcus dysgalactiae.
XX OS Streptococcus agalactiae.
XX OS Streptococcus parauberis.
XX OS Chimeric.
XX PN WO200196379-A2.
XX XX
XX XX 20-DEC-2001.

```

XX 11-JUN-2001; 2001WO-CA000836.
XX
XX
XX 12-JUN-2000; 2000US-0211247P.
XX
XX (UYSA-) UNIV SASKATCHEWAN.
XX
XX Potter AA, Perez-Casal J, Fontaine M;
XX
XX WPI; 2002-098051/13.
XX N-PSDB; ABA91327.
XX
XX Novel GapC multiple epitope fusion polypeptide comprising antigenic
XX determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
XX parauberis, or S. iniae GapC protein useful for treating mastitis in
XX vertebrates.
XX
XX Claim 9; Fig 6A-C; 116pp; English.
XX
XX The present sequence is that of a novel multiple epitope fusion protein,
XX designated PolyGap4, comprising the entire amino acid sequence of the
XX Streptococcus dysgalactiae GapC plasmin binding protein in addition to
XX unique amino acid sequences from the Streptococcus parauberis and
XX Streptococcus agalactiae GapC proteins. The multiple epitope protein is
XX produced in host cells transformed with an expression vector comprising a
XX chimeric gene (see ABA91327) encoding the protein. Polygal4 is an example
XX of novel GapC multiple epitope fusion proteins of the invention that
XX comprise epitopes from 1 or more of S. dysgalactiae, S. agalactiae, S.
XX parauberis, Streptococcus uberis and Streptococcus iniae (see AAM50665-
XX 69). The multiple epitope fusion proteins are used in claimed vaccines
XX for treating or preventing a bacterial infection in a vertebrate,
XX especially a streptococcal infection, and particularly mastitis. They are
XX also used in claimed methods of detecting Streptococcus antibodies. The
XX multiple epitope protein is capable of eliciting broad immunity against a
XX variety of streptococcal infections while minimising the number of
XX antigens present in the final formulation and concomitantly reducing
XX production costs. (Updated on 29-AUG-2003 to standardise OS field)
XX
XX Sequence 448 AA;
SQ

```

```

Query Match          96.6%; Score 1656.5; DB 5; Length 448;
Best Local Similarity 79.6%; Pred. No. 2.2e-139;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 1 MVYKVGINGFGRIGLAFRRIONVEGVETRLNDLTDPNMLAHLKYDTTQGRFDGTVEV 60
DB 28 MVYKVGINGFGRIGLAFRRIONVEGVETRLNDLTDPNMLAHLKYDTTQGRFDGTVEV 87
QY 61 KEGGFEVNGNFIKVSARDPE-----81
DB 88 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEALSGTVKXGDFVNGKFIKVSAAE 147
QY 82 -----NIDWATDGVETVLE 95
DB 148 KDPQIDWATDGVETVLEIDGTVEVKEGGFEVNGQFVKVSAREPENIDWATDGVETVLE 207
QY 96 ATGTFPAKKEAEKHLHANGAKKVVITAPGNDVKVTVFNTNHDILDTGTETVIGASCTTN 155
DB 208 ATSPFAKKEAEKHLHANGAKKVVITAPGNDVKVTVFNTNHDILDTGTETVIGASCTTN 267
QY 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAIVNPNSTGAA 215
DB 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAIVNPNSTGAA 327
QY 216 KAIGLVIPELNGKLDGAAQRPVPTGVSUTVELVTLDKNVSDVEINAAKASNDSPFGYTE 275
DB 328 KAIGLVIPELNGKLDGAAQRPVPTGVSUTVELVTLDKNVSDVEINAAKASNDSPFGYTE 387
QY 276 DPVSSDIVGVSGSLFDATQTKWMEVDGSQLVKVSVWYDNEMSYTAQLVTRTFYFAKIA 335
DB 388 DPVSSDIVGVSGSLFDATQTKWMEVDGSQLVKVSVWYDNEMSYTAQLVTRTFYFAKIA 447
QY 336 K 336

```

DB 448 K 448

RESULT 9

AAM50667
ID AAM50667 standard; protein; 336 AA.

XX AAM50667;

XX 08-APR-2002 (first entry)

XX Streptococcus uberis gapC plasmin binding protein UberGapC.

XX UberGapC; GapC; plasmin binding protein; epitope; infection; vaccine;

XX immunisation; mastitis; therapy.

XX Streptococcus uberis.

XX WO200196379-A2.

XX 20-DEC-2001.

XX 11-JUN-2001; 2001WO-CA000836.

XX 12-JUN-2000; 2000US-0211247P.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Potter AA, Perez-Casal J, Fontaine M;

XX WPI; 2002-098051/13.

XX N-PSDB; ABA91330.

XX Novel GapC multiple epitope fusion polypeptide comprising antigenic
XX determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
XX parauberis, or S. iniae GapC protein useful for treating mastitis in
XX vertebrates.

XX Claim 8; Fig 3A-B; 116pp; English.

XX The present sequence is that of the GapC plasmin binding protein,
XX UberGapC, of Streptococcus uberis ATCC 9527. The invention relates to
XX novel GapC multiple epitope fusion proteins that comprise epitopes from 1
XX or more of Streptococcus dysgalactiae, Streptococcus agalactiae,
XX Streptococcus parauberis, Streptococcus uberis and Streptococcus iniae
XX (see AAM50665-69). A claimed example is PolyGap4 (see AAM50664).
XX Expression vectors and host cells for production of the multiple epitope
XX fusion proteins are provided. The multiple epitope proteins are used in
XX claimed vaccines for treating or preventing a bacterial infection in a
XX vertebrate, especially a streptococcal infection, and particularly
XX mastitis. They are also used in claimed methods of detecting
XX Streptococcus antibodies. The multiple epitope proteins are capable of
XX eliciting broad immunity against a variety of streptococcal infections
XX while minimising the number of antigens present in the final formulation
XX and concomitantly reducing production costs

XX Sequence 336 AA;

Query Match 94.3%; Score 1618; DB 5; Length 336;

Best Local Similarity 92.9%; Pred. No. 4.1e-136;

Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

QY 1 MVYKVGINGFGRIGLAFRRIONVEGVETRLNDLTDPNMLAHLKYDTTQGRFDGTVEV 60

DB 1 MVYKVGINGFGRIGLAFRRIONVEGVETRLNDLTDPNMLAHLKYDTTQGRFDGTVEV 60

QY 61 KEGGFEVNGNFIKVSAREPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120

DB 61 KEGGFEVNGNFIKVSAREPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120

QY 121 TAPGNDVKTVFNTNHDILDTGTETVIGASCTTNCLAPMAKALHDAFGIQKGLMTTHIA 180

Db 121 TAPGGDDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALQDNFGVKQGLMTTIIHA 180
 QY 181 YTGQMLDGPGRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
 Db 181 YTGQMLDGPGRGDLRRARAGASNIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
 QY 241 GSVTELVTLDKNSVDEINAAKAAANDSFYTEDPIVSSDIVGVSYGSLFDATQTKVM 300
 Db 241 GSVTELVALEKETSVBEINAAKAAANDSYGYTEDPIVSSDIIGMAYGSLFDATQTKVQ 300
 QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
 Db 301 TVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 10
 AAM50641
 ID AAM50641 standard; protein; 336 AA.
 XX
 AC AAM50641;
 XX
 DT 04-APR-2002 (first entry)
 XX
 DE Streptococcus uberis gapC plasmin binding protein.
 XX
 KW GapC; plasmin-binding protein; UberGapC; infection; mastitis; vaccine;
 KW diagnosis; therapy.
 XX
 OS Streptococcus uberis.
 XX
 PN WO200196381-A2.
 XX
 PD 20-DEC-2001.
 XX
 PF 11-JUN-2001; 2001WO-CA000838.
 XX
 PR 12-JUN-2000; 2000US-0211022P.
 XX
 PA (UYSA-) UNIV SASKATCHEWAN.
 XX
 PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;
 XX
 DR WPI; 2002-130725/17.
 DR N-PSDB; ABA91250.
 XX
 DX Novel isolated GapC protein of Streptococcus dysgalactiae, S.agalactiae,
 PT S.uberis, S.paraberis, or S.iniae, useful as vaccine component for
 PT treating streptococcal infection which causes mastitis in vertebrates.
 XX
 PS Claim 1(c); Fig 3A-B; 107pp; English.
 XX
 CC The present sequence is that of the GapC plasmin binding protein
 CC (UberGapC) of Streptococcus uberis ATCC 9227, encoded by the GapC gene
 CC given in ABA91250. GapC, which has no signal sequence or membrane anchor
 CC domain, is capable of eliciting an immune response in a vertebrate. The
 CC invention provides the GapC genes and proteins of 5 Streptococcus
 CC species, as well as recombinant vectors, host cells and vaccine
 CC compositions comprising GapC polynucleotides or proteins. The vaccines
 CC are used to treat or prevent a bacterial infection, especially a
 CC streptococcal infection, and mastitis in particular (claimed). GapC
 CC proteins are also used in claimed methods for detecting GapC antibodies,
 CC and to raise antibodies that are used in claimed methods for detecting
 CC GapC proteins. S. uberis is a common pathogen associated with mastitis in
 CC cattle, horse, sheep and goat
 XX
 SQ Sequence 336 AA;
 CC
 CC Query Match 94.3%; Score 1618; DB 5; Length 336;
 CC Best Local Similarity 92.9%; Fred. No. 4.1e-136;
 CC Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;
 QY 1 MVTGKINGFGRIQRLAFERRIQNVGEVETRINDLTPNMLAHLKDYDTQGRFDGTVEV 60

Db 1 MVTGKINGFGRIQRLAFERRIQNVGEVETRINDLTPNMLAHLKDYDTQGRFDGTVEV 60
 QY 61 KEGGFEVNGNFVKVSAERDPENIDWATGVEIVLEATGFFFAKKEAAEKHLHANGAKKVI 120
 Db 61 KEGGFEVNGNFVKVSAERDPENIDWATGVEIVLEATGFFFAKKEAAEKHLHANGAKKVI 120
 QY 121 TAPGGNDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIIHA 180
 Db 121 TAPGGDDVKTVVNTNHDILDGTETVISGASCTTNCLAPMAKALQDNFGVKQGLMTTIIHA 180
 QY 181 YTGQMLDGPGRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
 Db 181 YTGQMLDGPGRGDLRRARAGASNIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
 QY 241 GSVTELVTLDKNSVDEINAAKAAANDSFYTEDPIVSSDIVGVSYGSLFDATQTKVM 300
 Db 241 GSVTELVALEKETSVBEINAAKAAANDSYGYTEDPIVSSDIIGMAYGSLFDATQTKVQ 300
 QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
 Db 301 TVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 11
 ABP30758
 ID ABP30758 standard; protein; 336 AA.
 XX
 AC ABP30758;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 10692.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX
 OS Streptococcus agalactiae.
 XX
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GB004789.
 XX
 PR 27-OCT-2000; 2000GB-00026333.
 PR 24-NOV-2000; 2000GB-00028727.
 PR 07-MAR-2001; 2001GB-00005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C;
 PI Tettelin H;
 XX
 DR WPI; 2002-352536/38.
 DR N-PSDB; ABN71389.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection or
 PT disease caused by Streptococcus bacteria, such as meningitis, and for
 PT detecting a compound that binds to the protein.
 XX
 PS Claim 1; Page 4179; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a

CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 XX
 SQ Sequence 336 AA;

Query Match 93.8%; Score 1609; DB 5; Length 336;
 Best Local Similarity 92.0%; Pred. No. 2.6e-135;
 Matches 309; Conservative 17; Mismatches 10; Indels 0; Gaps 0;
 QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETVTRINDLTDNMLAHLKYDTTQGRFDGTVEV 60
 DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETVTRINDLTDNMLAHLKYDTTQGRFDGTVEV 60
 QY 61 KEGGPEVNGNFIKVSAREDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
 DB 61 KEGGPEVNGQFVKVSAEREPANIDWATDGVETVLEATGFFAKKEAQEHENGAKKVI 120
 QY 121 TAPGGNDVKTVVNTNHDILDTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIIHA 180
 DB 121 TAPGGNDVKTVVNTNHDILDTETVISGASCTTNCLAPMAKALQDNFVKQGLMTTIIHA 180
 QY 181 YTGDMILDPHRRGGDLRRARAGANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
 DB 181 YTGDMILDPHRRGGDLRRARAGANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
 QY 241 GSVTELVTLTKNSVDENIAAMKAASNDSPGYTDEDPVSSDIVSGSLFDATQTKVM 300
 DB 241 GSVTELVATLEKDVTVVEVNAAMKAANDSYGTYDEDPVSSDIVSGSLFDATQTKVQ 300
 QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
 DB 301 TVDGNQLVKVSWYDNEMSYTSQLVRLTLEYFAKIAK 336

RESULT 12
 AAM50666
 ID AAM50666 standard; protein; 336 AA.
 AC AAM50666;
 XX
 DT 08-APR-2002 (first entry)
 XX
 DE Streptococcus agalactiae gapC plasmin binding protein AgalGapC.
 XX
 KW AgalGapC; GapC; plasmin binding protein; epitope; infection; vaccine;
 KW immunisation; mastitis; therapy.
 XX
 OS Streptococcus agalactiae.
 XX
 PN WO200196379-A2.
 XX
 PD 20-DEC-2001.
 XX
 PF 11-JUN-2001; 2001WO-CA000836.
 XX
 PR 12-JUN-2000; 2000US-0211247P.
 XX
 PA (UYSA-) UNIV SASKATCHEWAN.
 XX
 PI Potter AA, Perez-Casal J, Fontaine M;
 XX WPI; 2002-098051/13.
 DR N-PSDB; ABA91329.
 XX
 PT Novel GapC multiple epitope fusion polypeptide comprising antigenic
 PT determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
 PT parauberis, or S. iniae GapC protein useful for treating mastitis in

PT vertebrates.
 XX
 PS Claim 8; Fig 2A-B; 116pp; English.
 XX
 CC The present sequence is that of the gapC plasmin binding protein,
 CC AgalGapC, of Streptococcus agalactiae ATCC 27541. The invention relates
 CC to novel GapC multiple epitope fusion proteins that comprise epitopes
 CC from 1 or more of Streptococcus dysgalactiae, Streptococcus agalactiae,
 CC Streptococcus parauberis, Streptococcus uberis and Streptococcus iniae
 CC (see AAM50665-69). A claimed example is polyGap4 (see AAM50664).
 CC Expression vectors and host cells for production of the multiple epitope
 CC fusion proteins are provided. The multiple epitope proteins are used in
 CC claimed vaccines for treating or preventing a bacterial infection in a
 CC vertebrate, especially a streptococcal infection, and particularly
 CC mastitis. They are also used in claimed methods of detecting
 CC Streptococcus antibodies. The multiple epitope proteins are capable of
 CC eliciting broad immunity against a variety of streptococcal infections
 CC while minimising the number of antigens present in the final formulation
 CC and concomitantly reducing production costs
 XX
 SQ Sequence 336 AA;

Query Match 93.4%; Score 1602; DB 5; Length 336;
 Best Local Similarity 91.7%; Pred. No. 1.1e-134;
 Matches 308; Conservative 17; Mismatches 11; Indels 0; Gaps 0;
 QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETVTRINDLTDNMLAHLKYDTTQGRFDGTVEV 60
 DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETVTRINDLTDNMLAHLKYDTTQGRFDGTVEV 60
 QY 61 KEGGPEVNGNFIKVSAREDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
 DB 61 KEGGPEVNGQFVKVSAEREPANIDWATDGVETVLEATGFFAKKEAQEHENGAKKVI 120
 QY 121 TAPGGNDVKTVVNTNHDILDTETVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIIHA 180
 DB 121 TAPGGNDVKTVVNTNHDILDTETVISGASCTTNCLAPMAKALQDNFVKQGLMTTIIHA 180
 QY 181 YTGDMILDPHRRGGDLRRARAGANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
 DB 181 YTGDMILDPHRRGGDLRRARAGANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
 QY 241 GSVTELVTLTKNSVDENIAAMKAASNDSPGYTDEDPVSSDIVSGSLFDATQTKVM 300
 DB 241 GSVTELVATLEKDVTVVEVNAAMKAANDSYGTYDEDPVSSDIVSGSLFDATQTKVQ 300
 QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
 DB 301 TVDGNQLVKVSWYDNEMSYTSQLVRLTLEYFAKIAK 336

RESULT 13
 AAM50640
 ID AAM50640 standard; protein; 336 AA.
 AC AAM50640;
 XX
 DT 04-APR-2002 (first entry)
 XX
 DE Streptococcus agalactiae gapC plasmin binding protein.
 XX
 KW GapC; plasmin-binding protein; AgalGapC; infection; mastitis; vaccine;
 KW diagnosis; therapy.
 XX
 OS Streptococcus agalactiae.
 XX
 PN WO200196381-A2.
 XX
 PD 20-DEC-2001.
 XX
 PF 11-JUN-2001; 2001WO-CA000836.
 XX
 PR 12-JUN-2000; 2000US-0211022P.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 16:46:47 ; Search time 23 Seconds
(without alignments)
754.188 Million cell updates/sec

Title: US-10-650-369-12

Perfect score: 1715

Sequence: 1 MVVKVINGGRIGRLAFRR.....ENSYTAQLVLTLEYFAKIAK 336

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing printed 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1715	100.0	336	4	US-09-878-766A-12
2	1711	99.8	336	1	US-07-928-462-2
3	1711	99.8	336	3	US-08-273-247-2
4	1656.5	96.6	448	4	US-09-878-766A-22
5	1618	94.3	336	4	US-09-878-766A-16
6	1602	93.4	336	4	US-09-878-766A-14
7	1584	92.4	336	4	US-09-878-766A-20
8	1578	92.0	336	4	US-09-878-766A-18
9	1549.5	90.3	333	3	US-08-961-083-54
10	1549.5	90.3	333	4	US-09-536-784-54
11	1537.5	78.0	346	4	US-09-134-000C-4400
12	1325.5	77.3	333	4	US-09-107-532A-4769
13	1190	69.4	348	4	US-09-134-001C-5513
14	967	56.4	357	4	US-09-134-000C-4329
15	966	56.3	340	4	US-09-634-238-233
16	959	55.9	349	4	US-08-489-039A-13939
17	896.5	52.3	335	6	5290690-9
18	885.5	51.6	334	6	5290690-10
19	835.5	48.7	334	6	5290690-11
20	830	48.4	335	4	US-09-489-039A-7679
21	824.5	48.1	340	3	US-09-895-855-205
22	824.5	48.1	340	4	US-09-205-426-205
23	799	46.6	340	4	US-09-543-681A-8083
24	770.5	44.9	352	4	US-09-489-039A-9410
25	766.5	44.7	282	4	US-09-107-532A-4057
26	740.5	43.2	336	4	US-09-198-452A-662
27	736	42.9	338	4	US-09-091-725-10

ALIGNMENTS

RESULT 1

US-09-878-766A-12
; Sequence 12, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 12
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-09-878-766A-12

Query Match	100.0%	Score 1715;	DB 4;	Length 336;
Best Local Similarity	100.0%	Pred. No. 4.2e-164;		
Matches 336;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MVVKVINGGRIGRLAFRRIQNVGEVETRIINDLPNMLAHLKYDTTQGRFDGVEV	60	
Db	1	MVVKVINGGRIGRLAFRRIQNVGEVETRIINDLPNMLAHLKYDTTQGRFDGVEV	60	
Qy	61	KEGGEVNGNTIKVSAERDPENIDWATDGVIEATGFFAKKEAAEKHLHANGAKKVI	120	
Db	61	KEGGEVNGNTIKVSAERDPENIDWATDGVIEATGFFAKKEAAEKHLHANGAKKVI	120	
Qy	121	TAPGNDVKTVFTNTHDLDGTIVISGASCTTNCLAPMAKALHDAFGIOKGLMTTHA	180	
Db	121	TAPGNDVKTVFTNTHDLDGTIVISGASCTTNCLAPMAKALHDAFGIOKGLMTTHA	180	
Qy	181	YTGDMILDGPHRGDLRRRAGAAITVNSTGAAGAIGLVIPELNGKLDGAAORVPVPT	240	
Db	181	YTGDMILDGPHRGDLRRRAGAAITVNSTGAAGAIGLVIPELNGKLDGAAORVPVPT	240	
Qy	241	GSVTLVVTLTKNVSDVDEINAAKKAANDSPGYTDEPIVSSDIVGSYGSIFDQTQVM	300	
Db	241	GSVTLVVTLTKNVSDVDEINAAKKAANDSPGYTDEPIVSSDIVGSYGSIFDQTQVM	300	
Qy	301	EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK	336	
Db	301	EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK	336	

QY 301 EVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 336
 Db 301 EVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 3
 US-08-273-247-2
 ; Sequence 2, Application US/08273247
 ; Patent No. 6136323
 ; GENERAL INFORMATION:
 ; APPLICANT: Boyle, Michael D.P.
 ; APPLICANT: Lottenberg, Richard
 ; APPLICANT: Broder, Christopher C.
 ; APPLICANT: von Mering, Gregory O.
 ; TITLE OF INVENTION: Bacterial Plasmin Receptors as
 ; TITLE OF INVENTION: Fibrinolytic Agents
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: David R. Saliwanchik
 ; STREET: 2421 N.W. 41st Street, Suite A-1
 ; CITY: Gainesville
 ; STATE: FL
 ; COUNTRY: USA
 ; ZIP: 32606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/273,247
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/524,411
 ; FILING DATE: 16-MAY-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/330,849
 ; FILING DATE: 29-MAR-1989
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Saliwanchik, David R.
 ; REGISTRATION NUMBER: 31,794
 ; REFERENCE/DOCKET NUMBER: UF/S&S-13.C2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 904-375-8100
 ; TELEFAX: 904-372-5800
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 336 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-273-247-2

Query Match 99.8%; Score 1711; DB 3; Length 336;
 Best Local Similarity 99.4%; Pred. No. 1.1e-163;
 Matches 334; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVMKVGINGFGRIQLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
 Db 1 MVMKVGINGFGRIQLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60

QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIVLEATGFFAKKEAAEKHLHANGAKKVI 120
 Db 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIVLEATGFFAKKEAAEKHLHANGAKKVI 120

QY 121 TAPGGNDVTVVFNTHDILDTETVIGASCTTNC LAPMAKALHDAFGIQKGLMTTIIHA 180
 Db 121 TAPGGNDVTVVFNTHDILDTETVIGASCTTNC LAPMAKALHDAFGIQKGLMTTIIHA 180

QY 181 YTGDMQLDGPGRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
 Db 181 YTGDMQLDGPGRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240

US-07-928-462-2
 ; Sequence 2, Application US/07928462
 ; Patent No. 5328936
 ; GENERAL INFORMATION:
 ; APPLICANT: Boyle, Michael D.P.
 ; APPLICANT: Lottenberg, Richard
 ; APPLICANT: Broder, Christopher C.
 ; APPLICANT: von Mering, Gregory O.
 ; TITLE OF INVENTION: Bacterial Plasmin Receptors as
 ; TITLE OF INVENTION: Fibrinolytic Agents
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: David R. Saliwanchik
 ; STREET: 2421 N.W. 41st Street, Suite A-1
 ; CITY: Gainesville
 ; STATE: FL
 ; COUNTRY: USA
 ; ZIP: 32606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/928,462
 ; FILING DATE: 19920810
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/524,411
 ; FILING DATE: 16-MAY-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/330,849
 ; FILING DATE: 29-MAR-1989
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Saliwanchik, David R.
 ; REGISTRATION NUMBER: 31,794
 ; REFERENCE/DOCKET NUMBER: UF/S&S-13.C2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 904-375-8100
 ; TELEFAX: 904-372-5800
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 336 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-07-928-462-2

Query Match 99.8%; Score 1711; DB 1; Length 336;
 Best Local Similarity 99.4%; Pred. No. 1.1e-163;
 Matches 334; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVMKVGINGFGRIQLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
 Db 1 MVMKVGINGFGRIQLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60

QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIVLEATGFFAKKEAAEKHLHANGAKKVI 120
 Db 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIVLEATGFFAKKEAAEKHLHANGAKKVI 120

QY 121 TAPGGNDVTVVFNTHDILDTETVIGASCTTNC LAPMAKALHDAFGIQKGLMTTIIHA 180
 Db 121 TAPGGNDVTVVFNTHDILDTETVIGASCTTNC LAPMAKALHDAFGIQKGLMTTIIHA 180

QY 181 YTGDMQLDGPGRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
 Db 181 YTGDMQLDGPGRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240

QY 241 GSVTELVTLDKNVSDVINAMKASNDTSFGYTDPVSSDIVGVSGSLFDATQTKVM 300
 Db 241 GSVTELVTLDKNVSDVINAMKASNDTSFGYTDPVSSDIVGVSGSLFDATQTKVM 300

QY 241 GSVTELVVTLTKNVSVDENNAKAAASNDSPGVTEDPIVSSDIVGVSGLFDATQTKVM 300
DB 241 GSVTELVVTLTKNVSVDENNAKAAASNDSPGVTEDPIVSSDIVGVSGLFDATQTKVM 300
QY 301 EVDGSQLVKVSVYDNEYSYTAQLVRLTLEYFAKIAK 336
DB 301 EVDGSQLVKVSVYDNEYSYTAQLVRLTLEYFAKIAK 336

RESULT 4
US-09-878-766A-22
; Sequence 22, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878, 766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple
; OTHER INFORMATION: epitope fusion protein
US-09-878-766A-22

Query Match 96.6%; Score 1656.5; DB 4; Length 448;
Best Local Similarity 79.6%; Pred. No. 5.1e-158; Mismatches 1; Indels 85; Gaps 1;
Matches 335; Conservative 0

QY 1 MVKVGINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
DB 28 MVKVGINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 87
QY 61 KEGGFVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVI 81
DB 88 KEGGFVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVI 147
QY 82 -----NIDWATDGEIVLE 95
DB 148 KOPEQIDWATDGEIVLEIDGTVEVKEGGFVNGQFKVSAEREPANIDWATDGEIVLE 207
QY 96 ATGFFAKKEAAEKHLHANGAKKVIITAPCGNDVKTVFNTNHDILDGTETVISGASCTTN 155
DB 208 ATGFFAKKEAAEKHLHANGAKKVIITAPCGNDVKTVFNTNHDILDGTETVISGASCTTN 267
QY 156 CLAPMAKALHDAPGIGOKGLMTTHIATGQMLDGPGRGDLRRARAGANIVPNSTGAA 215
DB 268 CLAPMAKALHDAPGIGOKGLMTTHIATGQMLDGPGRGDLRRARAGANIVPNSTGAA 327
QY 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLTKNVSVDENNAKAAASNDSPGVT 275
DB 326 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLTKNVSVDENNAKAAASNDSPGVT 387
QY 276 DPTVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSVYDNEYSYTAQLVRLTLEYFAKIA 335
DB 388 DPTVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSVYDNEYSYTAQLVRLTLEYFAKIA 447
QY 336 K 336
DB 448 K 448

RESULT 5
US-09-878-766A-16

; Sequence 16, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878, 766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-09-878-766A-16

Query Match 94.3%; Score 1618; DB 4; Length 336;
Best Local Similarity 92.9%; Pred. No. 2.4e-154; Mismatches 9; Indels 0; Gaps 0;
Matches 312; Conservative 15

QY 1 MVKVGINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
DB 1 MVKVGINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
DB 61 KEGGFVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPCGNDVKTVFNTNHDILDGTETVISGASCTTNCLAPMAKALHDAPGIGOKGLMTTHI 180
DB 121 TAPCGNDVKTVFNTNHDILDGTETVISGASCTTNCLAPMAKALHDAPGIGOKGLMTTHI 180
QY 181 YTGQMLDGPGRGDLRRARAGANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
DB 181 YTGQMLDGPGRGDLRRARAGANIVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLTKNVSVDENNAKAAASNDSPGVTEDPIVSSDIVGVSGLFDATQTKVM 300
DB 241 GSVTELVVTLTKNVSVDENNAKAAASNDSPGVTEDPIVSSDIVGVSGLFDATQTKVM 300
QY 301 EVDGSQLVKVSVYDNEYSYTAQLVRLTLEYFAKIAK 336
DB 301 EVDGSQLVKVSVYDNEYSYTAQLVRLTLEYFAKIAK 336

RESULT 6
US-09-878-766A-14
; Sequence 14, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878, 766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-09-878-766A-14

Query Match 93.4%; Score 1602; DB 4; Length 336;
Best Local Similarity 91.7%; Pred. No. 9.9e-153; Mismatches 11; Indels 0; Gaps 0;
Matches 308; Conservative 17

QY 1 MVVKGINGFGRIGRLAFRRIONVEGEVETRIINDLTDNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKGINGFGRIGRLAFRRIONVEGEVETRIINDLTDNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFVNGNFIKVSAREDPENIDWATDGVIELEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGGFVNGNFIKVSAREDPENIDWATDGVIELEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVFNTNHDILDTETVTSAGSCTTNCNCLAPMAKALHDAFGIKQGLMTTIIA 180
Db 121 TAPGNDVKTIVFNTNHDILDTETVTSAGSCTTNCNCLAPMAKALQDNFVKGQGLMTTIIA 180
QY 181 YTGQMLDGHRRGDLRRRAGAAANVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGQMLDGHRRGDLRRRAGAAANVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVTLTKNVSVDENIINAAKAAANDSFGYTTEDPIVSSDIVGVSGLFDTATQTKV 300
Db 241 GSVTELVTLTKNVSVDENIINAAKAAANDSFGYTTEDPIVSSDIVGVSGLFDTATQTKV 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 TVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 7

US-09-878-766A-20
; Sequence 20, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 20
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus iniae
US-09-878-766A-20

Query Match 92.4%; Score 1584; DB 4; Length 336;
Best Local Similarity 90.5%; Pred. No. 6.4e-151;
Matches 304; Conservative 19; Mismatches 13; Indels 0; Gaps 0;

QY 1 MVVKGINGFGRIGRLAFRRIONVEGEVETRIINDLTDNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKGINGFGRIGRLAFRRIONVEGEVETRIINDLTDNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFVNGNFIKVSAREDPENIDWATDGVIELEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGGFVNGNFIKVSAREDPENIDWATDGVIELEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVFNTNHDILDTETVTSAGSCTTNCNCLAPMAKALHDAFGIKQGLMTTIIA 180
Db 121 TAPGNDVKTIVFNTNHDILDTETVTSAGSCTTNCNCLAPMAKALQDNFVKGQGLMTTIIA 180
QY 181 YTGQMLDGHRRGDLRRRAGAAANVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGQMLDGHRRGDLRRRAGAAANVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVTLTKNVSVDENIINAAKAAANDSFGYTTEDPIVSSDIVGVSGLFDTATQTKV 300
Db 241 GSVTELVTLTKNVSVDENIINAAKAAANDSFGYTTEDPIVSSDIVGVSGLFDTATQTKV 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 TVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 8

US-09-878-766A-18
; Sequence 18, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 18
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus parauberis
US-09-878-766A-18

Query Match 92.0%; Score 1578; DB 4; Length 336;
Best Local Similarity 90.8%; Pred. No. 2.6e-150;
Matches 305; Conservative 16; Mismatches 15; Indels 0; Gaps 0;

QY 1 MVVKGINGFGRIGRLAFRRIONVEGEVETRIINDLTDNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKGINGFGRIGRLAFRRIONVEGEVETRIINDLTDNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFVNGNFIKVSAREDPENIDWATDGVIELEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGGFVNGNFIKVSAREDPENIDWATDGVIELEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVFNTNHDILDTETVTSAGSCTTNCNCLAPMAKALHDAFGIKQGLMTTIIA 180
Db 121 TAPGNDVKTIVFNTNHDILDTETVTSAGSCTTNCNCLAPMAKALQDNFVKGQGLMTTIIA 180
QY 181 YTGQMLDGHRRGDLRRRAGAAANVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGQMLDGHRRGDLRRRAGAAANVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVTLTKNVSVDENIINAAKAAANDSFGYTTEDPIVSSDIVGVSGLFDTATQTKV 300
Db 241 GSVTELVTLTKNVSVDENIINAAKAAANDSFGYTTEDPIVSSDIVGVSGLFDTATQTKV 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 TVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 9

US-08-961-083-54
; Sequence 54, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961.083
FILING DATE: 30-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961.083
FILING DATE: OCT-30-1997
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-08-961-083-54

Query Match 90.3%; Score 1549.5; DB 3; Length 333;
Best Local Similarity 91.3%; Pred. No. 1.9e-147;
Matches 303; Conservative 12; Mismatches 16; Indels 1; Gaps 1;
QY 2 VVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVYVK 61
DB 1 VVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVYVK 60
QY 62 EGGFEVNGNFKVSAERDPENIDWATGVEIVLEATGFFAKKEAAEKHLHANGAKKVVIT 121
DB 61 EGGFEVNGKFKVSAERDPQIDWATGVEIVLEATGFFAKKEAAEKHL-KGGAKKVVIT 119
QY 122 APGNDVKTVPNTNHDLDGTETVISCATTNCLAPMAKALHDAPGIQKGLMTTHAY 181
DB 120 APGNDVKTVPNTNHDLDGTETVISCATTNCLAPMAKALQDNFVGVEGLMTTHAY 179
QY 182 TGDQWILDGPHRGDLRRARAGANIVPNTGAAKALGLVPELNGKLDGAAQRPVPTG 241
DB 180 TGDQWILDGPHRGDLRRARAGANIVPNTGAAKALGLVPELNGKLDGSAQRPVPTG 239
QY 242 SVTELAVTLKXNSVDEINAAKASNDSPGYTDEPIVSSDIVGVSGLFDATQTKVME 301
DB 240 SVTELAVLEKNVTVDENAAKASNESGYTDEPIVSSDIVGMSYGLFDATQTKVLD 299
QY 302 VDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAK 333
DB 300 VDGKQLVKVSWYDNEMSYTAQLVRLTGLILRK 331

RESULT 10

US-09-536-784-54

Sequence 54, Application US/09536784

Patent No. 6573082

GENERAL INFORMATION:

APPLICANT: Choi et. al.

TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines

NUMBER OF SEQUENCES: 452

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/536,784

FILING DATE: 30-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961.083
FILING DATE: OCT-30-1997
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB340P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-09-536-784-54

Query Match 90.3%; Score 1549.5; DB 4; Length 333;
Best Local Similarity 91.3%; Pred. No. 1.9e-147;
Matches 303; Conservative 12; Mismatches 16; Indels 1; Gaps 1;
QY 2 VVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVYVK 61
DB 1 VVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVYVK 60
QY 62 EGGFEVNGNFKVSAERDPENIDWATGVEIVLEATGFFAKKEAAEKHLHANGAKKVVIT 121
DB 61 EGGFEVNGKFKVSAERDPQIDWATGVEIVLEATGFFAKKEAAEKHL-KGGAKKVVIT 119
QY 122 APGNDVKTVPNTNHDLDGTETVISCATTNCLAPMAKALHDAPGIQKGLMTTHAY 181
DB 120 APGNDVKTVPNTNHDLDGTETVISCATTNCLAPMAKALQDNFVGVEGLMTTHAY 179
QY 182 TGDQWILDGPHRGDLRRARAGANIVPNTGAAKALGLVPELNGKLDGAAQRPVPTG 241
DB 180 TGDQWILDGPHRGDLRRARAGANIVPNTGAAKALGLVPELNGKLDGSAQRPVPTG 239
QY 242 SVTELAVTLKXNSVDEINAAKASNDSPGYTDEPIVSSDIVGVSGLFDATQTKVME 301
DB 240 SVTELAVLEKNVTVDENAAKASNESGYTDEPIVSSDIVGMSYGLFDATQTKVLD 299
QY 302 VDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAK 333
DB 300 VDGKQLVKVSWYDNEMSYTAQLVRLTGLILRK 331

RESULT 11

US-09-134-000C-4400

Sequence 4400, Application US/09134000C

Patent No. 6617156

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

NUMBER OF SEQUENCES: 452

FILE REFERENCE: 032796-032

CURRENT APPLICATION NUMBER: US/09/134,000C

CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/055,778

PRIOR FILING DATE: 1997-08-15

NUMBER OF SEQ ID NOS: 6812

SOFTWARE: PatentIn version 3.1

SEQ ID NO 4400

LENGTH: 346

TYPE: PRT

ORGANISM: Enterococcus faecalis

US-09-134-000C-4400

Query Match 78.0%; Score 1337.5; DB 4; Length 346;

Best Local Similarity 78.4%; Pred. No. 4.1e-126; Mismatches 24; Indels 1; Gaps 1;
Matches 262; Conservative 24; Mismatches 47; Indels 1; Gaps 1;
QY 1 MVVKVINGPGRIGRLAFRRIONVEGVETVRINDLTDPMMLAHLKYDTTQGRPDGTVEV 60
Db 14 MTVKVINGPGRIGRLAFRRIQDVEGVETVRINDLTDKMLAHLKYDTTQGRPDGTVEV 73
QY 61 KEGGFVNGNFKVSAERDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 74 HEGSFVNGKEIKVLANRNPPELPGELGVDIVLECTGFTTSCAAEKHLTA-GAKRVVI 132
QY 121 TAPGNDVKTIVFNNHDLIDGTETVSGASCTTNCCLAPMAKALHDAFGIOKGLMTTIHA 180
Db 133 SAPGNDVPTIVYNTNHTLAGEETVSGASCTTNCCLAPMAKALHDAFGIOKGLMTTIHA 192
QY 181 YTGDMILDGHRRGDLRRARAGANIVPNSGAAKALGLVPELNGKLDGAAQRPVPT 240
Db 193 YTGDMITLDGHPKGDFFRARAANIVPNSGAAKALGLVPELNGKLDGAAQRPVPT 252
QY 241 GSVTELVVTLDKNVSVDENIAMAASNDSTGYTETDPIVSSDIVSGSLFDATQTKVM 300
Db 253 GSVTELVVTLDEKVTVDENIAMAASNDSTGYTETDPIVSSDIVSGSLFDATQTKVM 312
QY 301 EVDGSOLVKVSWYDNEMSYTAQLVRLTLEYFAKI 334
Db 313 TVGDKQLVKTVAWYDNEMSYTAQLVRLTLEYFANL 346
RESULT 12
US-09-107-532A-4769
; Sequence 4769, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Denete
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4769:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 333 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium

FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...333
; SEQUENCE DESCRIPTION: SEQ ID NO: 4769:
US-09-107-532A-4769
Query Match 77.3%; Score 1325.5; DB 4; Length 333;
Best Local Similarity 77.8%; Pred. No. 6.3e-125;
Matches 260; Conservative 25; Mismatches 48; Indels 1; Gaps 1;
QY 1 MVVKVINGPGRIGRLAFRRIONVEGVETVRINDLTDPMMLAHLKYDTTQGRPDGTVEV 60
Db 1 MTVKVINGPGRIGRLAFRRIQDVEGVETVRINDLTDKMLAHLKYDTTQGRPDGTVEV 60
QY 61 KEGGFVNGNFKVSAERDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 HEGSFVNGKEIKVLANRNPPELPGELGVDIVLECTGFTTSCAAEKHLTA-GAKRVVI 119
QY 121 TAPGNDVKTIVFNNHDLIDGTETVSGASCTTNCCLAPMAKALHDAFGIOKGLMTTIHA 180
Db 120 SAPGNDVPTIVYNTNHTLAGEETVSGASCTTNCCLAPMAKALHDAFGIOKGLMTTIHA 179
QY 181 YTGDMILDGHRRGDLRRARAGANIVPNSGAAKALGLVPELNGKLDGAAQRPVPT 240
Db 180 YTGDMITLDGHPKGDFFRARAANIVPNSGAAKALGLVPELNGKLDGAAQRPVPT 239
QY 241 GSVTELVVTLDKNVSVDENIAMAASNDSTGYTETDPIVSSDIVSGSLFDATQTKVM 300
Db 240 GSVTELVVTLDEKVTVDENIAMAASNDSTGYTETDPIVSSDIVSGSLFDATQTKVM 299
QY 301 EVDGSOLVKVSWYDNEMSYTAQLVRLTLEYFAKI 334
Db 300 TVGDKQLVKTVAWYDNEMSYTAQLVRLTLEYFANL 333
RESULT 13
US-09-134-001C-5513
; Sequence 5513, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5513
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5513
Query Match 69.4%; Score 1190; DB 4; Length 348;
Best Local Similarity 69.4%; Pred. No. 2.9e-111;
Matches 234; Conservative 39; Mismatches 62; Indels 2; Gaps 2;
QY 1 MVVKVINGPGRIGRLAFRRIONVEGVETVRINDLTDPMMLAHLKYDTTQGRPDGTVEV 60
Db 13 MAIKVAINGPGRIGRLAFRRIQDVEGVETVRINDLTDKMLAHLKYDTTQGRPDGTVEV 72
QY 61 KEGGFVNGNFKVSAERDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 73 IEGGFVNGKEIKVLPDAGKLPWGDLDDIVLECTGFTTSCAAEKHLTA-GAKKVI 131
QY 121 TAPGNDVKTIVFNNHDLIDGTETVSGASCTTNCCLAPMAKALHDAFGIOKGLMTTIHA 180
Db 132 SAPAKGDVKTIVFNNHDLIDGSETVYSGASCTTNSLAPVAKVLSDFGLVEGFMPTTIHA 191

QY 181 YTGDMILDPHRRGDLRRRAGAAANIVPNSGTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
 Db 192 YTGDMTQDAPHRRGDKRRARAENIIPNSTGAAGAIGLVIPEIDGKLDGGAAQRPVPT 251
 QY 241 GSVTELVVTLTK-NVSVDEINAAKASNDSPGYTEDPIVSSDIVGVSGLSDATQTKV 299
 Db 252 GSLTELVVTLTKQVTVQVNSAMKQASDSFGYTEDEIVSSDIVGVSGLSDATQTRV 311
 QY 300 MEYDGSOLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
 Db 312 MTVDROLVQVAWYDNEMSYTAQLVRLTLEYFAKIAK 348

RESULT 14
 US-09-134-000C-4229
 ; Sequence 4229, Application US/09134000C
 ; Patent No. 6617156
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; FILE REFERENCE: 032796-032
 ; CURRENT APPLICATION NUMBER: US/09/134.000C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/055,778
 ; PRIOR FILING DATE: 1997-08-15
 ; NUMBER OF SEQ ID NOS: 6812
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4229
 ; LENGTH: 357
 ; TYPE: PRT
 ; ORGANISM: Enterococcus faecalis
 US-09-134-000C-4229

Query Match 56.4%; Score 967; DB 4; Length 357;
 Best Local Similarity 58.3%; Pred. No. 8e-89;
 Matches 197; Conservative 44; Mismatches 91; Indels 6; Gaps 5;

QY 1 MVVKVINGGFRIGRLAPRIQNV-EGVEVTRINDLTDPNMLAHLKYDTTQGRFDGTV 59
 Db 22 MTVKVINGGFRIGRLAPRIQNV-EGVEVTRINDLTDPNMLAHLKYDTTQGRFDGTV 81
 QY 60 VKEGFFVNGNFIKVSARDPENIDWATD-GVIVLEATGFFAKKEAAEKHLHANGAKV 118
 Db 82 ATENGIIVDGEERVRVAPASKIPWVKENGVDIVLECTGYTSEKAQAHLDA-GVKRV 140
 QY 119 VITAPGNDVKTVFNTHDILDTETVTSAGCTTNCLAPMAKALHDAFGIKGLMTTI 178
 Db 141 VISAPAG-AMKTIYVNVDDTLDAKDIIISAGCTTNCLAPMAYFLNNEFIEVGTMTTV 199
 QY 179 HAYTGDQMLDGPVRRGDLRRRAGAAANIVPNSGTGAAGAIGLVIPELNGKLDGAAQRPV 238
 Db 200 HAYTSTQMLDGPVRRGDLRRRAGAAANIVPNSGTGAAGAIGLVIPELNGKLDGAAQRPV 259
 QY 239 PTGSVTELVVTLTKNVSVDENIAAMK--AASNDSPGYTEDPIVSSDIVGVSGLSDATQ 296
 Db 260 VDGSLELVSLTKVKTADQVNEAMKHTIDNPSFGYDDREIVSGDIIGTTEGSIFFDPTQ 319
 QY 297 TKMVEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 334
 Db 320 TEVTTAGDQVLVKTVAWYDNEMSYTAQLVRLTLEYFAKIAK 357

RESULT 15
 US-09-634-238-233
 ; Sequence 233, Application US/09634238
 ; Patent No. 654772
 ; GENERAL INFORMATION:
 ; APPLICANT: Glenn, Matthew
 ; APPLICANT: Havukkala, Ilkka J.
 ; APPLICANT: Bloksberg, Leonard, N.
 ; APPLICANT: Lubbers, Mark W.
 ; APPLICANT: Dekker, James

; APPLICANT: Christensson, Anna C.
 ; APPLICANT: Holland, Ross
 ; APPLICANT: O'Toole, Paul W.
 ; APPLICANT: Reid, Julian R.
 ; TITLE OF INVENTION: Polynucleotides, materials incorporating
 ; TITLE OF INVENTION: them and methods for using them.
 ; FILE REFERENCE: 11000.1043UI
 ; CURRENT APPLICATION NUMBER: US/09/634.238
 ; CURRENT FILING DATE: 2000-08-08
 ; NUMBER OF SEQ ID NOS: 422
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 233
 ; LENGTH: 340
 ; TYPE: PRT
 ; ORGANISM: Lactobacillus rhamnosus
 US-09-634-238-233

Query Match 56.3%; Score 966; DB 4; Length 340;
 Best Local Similarity 57.0%; Pred. No. 9.4e-89;
 Matches 195; Conservative 48; Mismatches 89; Indels 10; Gaps 6;

QY 1 MVVKVINGGFRIGRLAPRIQNV-EGVEVTRINDLTDPNMLAHLKYDTTQGRFDG 56
 Db 1 MTVKVINGGFRIGRLAPRIQNV-EGVEVTRINDLTDPNMLAHLKYDTTQGRFDG 60
 QY 57 TVEVKEGFFVNGNFIKVSARDPENIDWA-TDGVIVLEATGFFAKKEAAEKHLHANGA 115
 Db 61 EVSATDNGIIVDGEERVRVAPASKIPWVKENGVDIVLECTGYTSEKAQAHLDA-GA 119
 QY 116 KKWVITAPGNDVKTVFNTHDILDTETVTSAGCTTNCLAPMAKALHDAFGIKGLM 175
 Db 120 KRVLSAPAGK-IKTIYVNVDDTLNADDKIVSAGCTTNCLAPMAYFLNNEFIEVGTMT 178
 QY 176 TTIHAYTGDQMLDGPVRRGDLRRRAGAAANIVPNSGTGAAGAIGLVIPELNGKLDGAAQ 235
 Db 179 TTVHAYTSTQMLDGPVRRGDLRRRAGAAANIVPNSGTGAAGAIGLVIPELNGKLDGAAQ 238
 QY 236 VPVPTGSVTELVVTLTKNVSVDENIAAMK--AASNDSPGYTEDPIVSSDIVGVSGLS 292
 Db 239 VSVVDGSLTELVSLTKVKTADQVNEAMKHTIDNPSFGYDDREIVSGDIIGTTEGSI 298
 QY 293 DATQTKMVEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 334
 Db 299 DPTQTEVTTAGDQVLVKTVAWYDNEMSYTAQLVRLTLEYFAKIAK 340

Search completed: May 11, 2004, 16:50:40
 Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 16:45:46 ; Search time 21 Seconds
(without alignments)
1539.064 Million cell updates/sec

Title: US-10-650-369-12
Perfect score: 1715
Sequence: 1 MVKVGINGFGRIGRLAFRR.....EMSYTQLVLTLEYPFAKIAK 336

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1711	99.8	336	2 A42963	glyceraldenhyde-3-p
2	1617	94.3	336	2 S71350	glyceraldenhyde-3-p
3	1593.5	92.9	335	2 P95235	glyceraldenhyde 3-p
4	1593.5	92.9	359	2 G98039	glyceraldenhyde-3-p
5	1410	82.2	336	2 P86905	hypothetical prote
6	1338.5	78.0	337	2 G86634	hypothetical prote
7	1289.5	75.2	334	2 S34254	glyceraldenhyde-3-p
8	1262	73.6	336	2 ACl382	glyceraldenhyde 3-p
9	1261	73.5	336	2 AD1751	glyceraldenhyde 3-p
10	1259.5	73.4	334	2 C96987	glyceraldenhyde 3-p
11	1197.5	69.8	334	2 B82019	glyceraldenhyde-3-p
12	1184.5	69.1	334	2 E81001	glyceraldenhyde 3-p
13	1183	69.0	336	2 E89850	glyceraldenhyde-3-p
14	1003	58.5	338	2 T09633	glyceraldenhyde-3-p
15	993	57.9	333	2 F90881	glyceraldenhyde 3-p
16	993	57.9	333	2 C85737	glyceraldenhyde-3-p
17	921	53.7	337	2 S73737	glyceraldenhyde-3-p
18	909.5	53.0	336	2 A43260	glyceraldenhyde-3-p
19	908	52.9	337	2 C64233	glyceraldenhyde-3-p
20	907	52.9	349	2 F90517	glyceraldenhyde 3-p
21	896.5	52.3	335	1 DBSG	glyceraldenhyde 3-p
22	891.5	52.0	335	2 S12696	glyceraldenhyde-3-p
23	885.5	51.6	335	1 DBSGF	glyceraldenhyde 3-p
24	882.5	51.5	335	2 H84034	glyceraldenhyde-3-p
25	874	51.0	333	1 DBHGT	glyceraldenhyde-3-p
26	863	50.3	335	2 A70107	probable glycerald
27	843	49.2	342	2 F70391	glyceraldenhyde-3-p
28	842.5	49.1	336	2 T36020	glyceraldenhyde-3-p
29	831	48.5	334	2 A10262	glyceraldenhyde-3-p

ALIGNMENTS

RESULT 1

A42963
glyceraldenhyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Streptococcus N; Alternate names: plasmin receptor
C: Species: Streptococcus sp.
C: Date: 10-Mar-1994 #sequence revision 10-Mar-1994 #text_change 03-Jun-2002
C: Accession: A42963; B42963; JH0750
R: Lottenberg, R.; Broder, C.C.; Boyle, M.D.; Kain, S.J.; Schroeder, B.L.; Curtiss III, F. J. Bacteriol. 174, 5204-5210, 1992
A: Title: Cloning, sequence analysis, and expression in Escherichia coli of a streptococcal reference number: A42963; MUID:92355491; PMID:1322883
A: Accession: A42963
A: Molecule type: DNA
A: Residues: 1-336 <LOT>
A: Experimental source: Group A, strain 64/14
A: Note: sequence extracted from NCBI backbone (NCBIP:110308)
A: Accession: B42963
A: Molecule type: protein
A: Residues: 2-74; 161-164, 'X', 166-174; 187-211, 'X', 213-217 <LO2>
R: Pancholi, V.; Fischetti, V.A.
J. Exp. Med. 176, 415-426, 1992
A: Title: A major surface protein on group A streptococci is a glyceraldehyde-3-phosphate dehydrogenase
A: Reference number: JH0750; MUID:92364544; PMID:1500854
A: Accession: JH0750
A: Molecule type: protein
A: Residues: 2-30, 'A', 32-40 <PAN>
C: Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C: Keywords: gluconogenesis; glycolysis; homotetramer; NAD; oxidoreductase
F: 152/Active site: Cys #status predicted

Query Match 99.8%; Score 1711; DB 2; Length 336;
Best Local Similarity 99.4%; Pred. No. 4.2e-116;
Matches 334; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY	1	MVKVGINGFGRIGRLAFRRIONVEGVETRI	INDLTDENMLAHLAKYDTTQGRFGTVEV	60
DB	1	MVKVGINGFGRIGRLAFRRIONVEGVETRI	INDLTDENMLAHLAKYDTTQGRFGTVEV	60
QY	61	KEGGFEVNGNFVKVSAERDPENIDWATGVEI	LEATGFFAKKEAAEKHLHANGAKKVI	120
DB	61	KEGGFEVNGNFVKVSAERDPENIDWATGVEI	LEATGFFAKKEAAEKHLHANGAKKVI	120
QY	121	TAPGGNDVKTVPVNTNHDLDGTETVIGASCTT	NCLAPMAKALHDAFGIQKGLMTTTHA	180
DB	121	TAPGGNDVKTVPVNTNHDLDGTETVIGASCTT	NCLAPMAKALHDAFGIQKGLMTTTHA	180
QY	181	YTGDMQILDGPHRGDLRRARAGANIVPNSGAA	KAIGLVIPELNGLDGAAPVPPT	240
DB	181	YTGDMQILDGPHRGDLRRARAGANIVPNSGAA	KAIGLVIPELNGLDGAAPVPPT	240
QY	241	GSVTLEVTLDKNSVDEINAMKAASNDSPGTT	DPVSSDIVGSVGSIFDATQTKM	300
DB	241	GSVTLEVTLDKNSVDEINAMKAASNDSPGTT	DPVSSDIVGSVGSIFDATQTKM	300

Db 241 GSVTELVVTLTKNVSVDEINAMKAASNDSPGYTDPVSSDIVGVSGSLFDATQTKVM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 2
S71350
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Streptococcus
C:Species: Streptococcus "equisimilis"
C:Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 03-Jun-2002
C:Accession: S71350
R:Gase, K.; Gase, A.; Schirmer, H.; Malke, H.
Eur. J. Biochem. 239, 42-51, 1996
A:Title: Cloning, sequencing and functional overexpression of the Streptococcus equisimilis
inducing protein. Purification and biochemical characterization of the protein.
A:Reference number: S71350; MUID:96303364; PMID:8706717
A:Accession: S71350
A:Molecule type: DNA
A:Residues: 1-336 <GAS>
A:Cross-references: EMBL:X97788; NID:g1478268; PIDN:CAA66377.1; PID:g1478269
A:Experimental source: strain H46A
C:Genetics:
A:Gene: GapC
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C:Keywords: gluconeogenesis; glycolysis; NAD; oxidoreductase
F:4-34/Region: beta-alpha-beta NAD nucleotide-binding fold
F:152/Active site: Cys #status predicted

Query Match 94.3%; Score 1617; DB 2; Length 336;
Best Local Similarity 95.5%; Pred. No. 2.5e-109; Indels 0; Gaps 0;
Matches 321; Conservative 2; Mismatches 13;

QY 1 MVVKVINGFGFRIGRLAFRRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGFRIGRLAFRRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60

QY 61 KEGGFVNGNFIKVSARDPENIDWATGVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGGFVNGNFIKVSARDPENIDWATGVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120

QY 121 TAPGGNDVKTVVFNTHDILDTETVIGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180
Db 121 TAPGGNDVKTVVFNTHDILDTETVIGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180

QY 181 YTGDMILDPHRRGDLRRRAGAANIIPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDPHRRGDLRRRAGAANIIPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240

QY 241 GSVTELVVTLTKNVSVDEINAMKAASNDSPGYTDPVSSDIVGVSGSLFDATQTKVM 300
Db 241 GSVTELVVTLTKNVSVDEINAMKAASNDSPGYTDPVSSDIVGVSGSLFDATQTKVM 300

QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 3
F95235
glyceraldehyde 3-phosphate dehydrogenase [imported] - Streptococcus pneumoniae (strain T
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C:Accession: F95235
R:Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.C.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: F95235

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-335 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK76079.1; PID:g14973522; GSPDB:GN00164; TIGR:SP;
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP2012
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 92.9%; Score 1593.5; DB 2; Length 335;
Best Local Similarity 92.6%; Pred. No. 1.1e-107; Indels 1; Gaps 1;
Matches 311; Conservative 12; Mismatches 12;

QY 1 MVVKVINGFGFRIGRLAFRRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGFRIGRLAFRRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60

QY 61 KEGGFVNGNFIKVSARDPENIDWATGVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGGFVNGNFIKVSARDPENIDWATGVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120

QY 121 TAPGGNDVKTVVFNTHDILDTETVIGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180
Db 121 TAPGGNDVKTVVFNTHDILDTETVIGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180

QY 181 YTGDMILDPHRRGDLRRRAGAANIIPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDPHRRGDLRRRAGAANIIPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240

QY 241 GSVTELVVTLTKNVSVDEINAMKAASNDSPGYTDPVSSDIVGVSGSLFDATQTKVM 300
Db 241 GSVTELVVTLTKNVSVDEINAMKAASNDSPGYTDPVSSDIVGVSGSLFDATQTKVM 300

QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 4
G98099
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) [imported] - S:
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 03-Jun-2002
C:Accession: G98099
R:Hoekings, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; E
e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; N
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: G98099
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-359 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAL00628.1; PID:g15459513; GSPDB:GN00174
C:Genetics:
A:Gene: GapA
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C:Keywords: oxidoreductase

Query Match 92.9%; Score 1593.5; DB 2; Length 359;
Best Local Similarity 92.6%; Pred. No. 1.4e-107; Indels 1; Gaps 1;
Matches 311; Conservative 12; Mismatches 12;

QY 1 MVVKVINGFGFRIGRLAFRRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
Db 25 MVVKVINGFGFRIGRLAFRRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 84

QY 61 KEGGFVNGNFIKVSARDPENIDWATGVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 85 KEGGFVNGNFIKVSARDPENIDWATGVEIVLEATGFFAKKEAAEKHLHANGAKKVI 143

QY 121 TAPGNDVKTIVFNTNHDLDGTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180
Db 144 TAPGNDVKTIVFNTNHDLDGTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 203
QY 181 YTGQMDLDPHRRGDLRRARAGAANTVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 204 YTGQMDLDPHRRGDLRRARAGAANTVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 263
QY 241 GSVTELVVTLTKNVSDVEINAAKAAASNDSPGYTETDPIVSSDIVGVSGLFATQTKVM 300
Db 264 GSVTELVVTLTKNVSDVEINAAKAAASNDSPGYTETDPIVSSDIVGVSGLFATQTKVL 323
QY 301 EVDGSQLVKVSVYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 324 DVDGSQLVKVSVYDNEMSYTAQLVRLTLEYFAKIAK 359
RESULT 5
F86905
hypothetical protein gapB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: F86905
R:Boletín, A.; Wincker, P.; Møgel, S.; Jaillon, O.; Malmgren, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: F86905
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-336 <STO>
A:Cross-references: GB:AE005176; PID:g12725315; PIDN:AAK06344.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: gapB
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
Query Match 82.2%; Score 1410; DB 2; Length 336;
Best Local Similarity 81.2%; Pred. No. 2.2e-94;
Matches 273; Conservative 22; Mismatches 41; Indels 0; Gaps 0;
QY 1 MVVKVINGFGRIQGLAFRRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVVEV 60
Db 1 MVVKVINGFGRIQGLAFRRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVVEV 60
QY 61 KEGGFVNGNFIKVSARDPENIDWATGVEIVLEATGFFAKKEAEKHLHANGAKKVI 120
Db 61 KEGGFVNGNFIKVSARDPENIDWATGVEIVLEATGFFAKKEAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVFNTNHDLDGTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180
Db 121 TAPGNDVKTIVFNTNHDLDGTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180
QY 181 YTGQMDLDPHRRGDLRRARAGAANTVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGQMDLDPHRRGDLRRARAGAANTVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLTKNVSDVEINAAKAAASNDSPGYTETDPIVSSDIVGVSGLFATQTKVM 300
Db 241 GSVTELVVTLTKNVSDVEINAAKAAASNDSPGYTETDPIVSSDIVGVSGLFATQTKVM 300
QY 301 EVDGSQLVKVSVYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 SADGSQLVKVSVYDNEMSYTNLVRITLAVFAKIAK 336
RESULT 6
G86694
hypothetical protein gapA [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: G86694
R:Boletín, A.; Wincker, P.; Møgel, S.; Jaillon, O.; Malmgren, K.; Weissenbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: G86694
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <STO>
A:Cross-references: GB:AE005176; PID:g12723446; PIDN:AAK04657.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: gapA
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
Query Match 78.0%; Score 1338.5; DB 2; Length 337;
Best Local Similarity 77.2%; Pred. No. 3.2e-89;
Matches 260; Conservative 31; Mismatches 45; Indels 1; Gaps 1;
QY 1 MVVKVINGFGRIQGLAFRRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVVEV 60
Db 1 MVVKVINGFGRIQGLAFRRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVVEV 60
QY 61 KEGGFVNGNFIKVSARDPENIDWATGVEIVLEATGFFAKKEAEKHLHANGAKKVI 120
Db 61 KEGGFVNGNFIKVSARDPENIDWATGVEIVLEATGFFAKKEAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVFNTNHDLDGTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180
Db 121 TAPGNDVKTIVFNTNHDLDGTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180
QY 181 YTGQMDLDPHRRGDLRRARAGAANTVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGQMDLDPHRRGDLRRARAGAANTVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLTKNVSDVEINAAKAAASNDSPGYTETDPIVSSDIVGVSGLFATQTKVM 300
Db 241 GSVTELVVTLTKNVSDVEINAAKAAASNDSPGYTETDPIVSSDIVGVSGLFATQTKVM 300
QY 301 EVDGSQLVKVSVYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 DLKDGSQLVKVSVYDNEMSYTAQLVRLTLEYFAKIAK 337
RESULT 7
S34254
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Clostridium i
C:Species: Clostridium pasteurianum
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 03-Jun-2002
C:Accession: S34254
R:Oster, T.; Assobhei, O.; Scherrer, S.; Branlant, G.; Branlant, C.
submitted to the EMBL Data Library, May 1993
A:Description: Nucleotide sequence of the glyceraldehyde-3-phosphate dehydrogenase
A:Reference number: S34254
A:Accession: S34254
A:Molecule type: DNA
A:Residues: 1-334 <OST>
A:Cross-references: EMBL:X72219; NID:g311923; PIDN:CAA51020.1; PID:g311924
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C:Keywords: oxidoreductase
Query Match 75.2%; Score 1289.5; DB 2; Length 334;
Best Local Similarity 74.3%; Pred. No. 1.1e-85;
Matches 249; Conservative 33; Mismatches 52; Indels 1; Gaps 1;
QY 2 VVKVINGFGRIQGLAFRRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVVEV 61
Db 1 MTKVAINGFGRIQGLAFRRIONVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVVEV 60
QY 62 EGGFVNGNFIKVSARDPENIDWATGVEIVLEATGFFAKKEAEKHLHANGAKKVI 121
Db 61 EGAFVNGNFIKVSARDPENIDWATGVEIVLEATGFFAKKEAEKHLHANGAKKVI 119
QY 122 APGNDVKTIVFNTNHDLDGTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 181
Db 122 APGNDVKTIVFNTNHDLDGTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 181

Db 120 APAGNDLKTIVFNANEDLDGTETVISGASCTTNCCLAPMAKVLNDKFGIEKGFMTTHAY 179
 QY 182 TGDQMLDGPGRGDLRRAGAANIVPNSTGAAGAIGLVIPELNGKLDGAARVPVPTG 241
 Db 180 TNDQNTLDGPHRGKDFRRARAASIIENSTGAAKAIAQVPELKGKLDGNAQRPVPTG 239
 QY 242 SVTELVTLDKXVSDVDEINAAKASNDGFGTEPDIYSSDIVGVSGSLFDATQTKME 301
 Db 240 SVTELISVJKNVTVEEINAAKESFGTEDEIVSADVGVGSLGSLFDATLTKIVD 299
 QY 302 VDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
 Db 300 VDGSQLVKVSWYDNEMSYTSLVRLTLEYFAKIAK 334

RESULT 8
 AC1382
 Glyceraldehyde 3-phosphate dehydrogenase homolog gap [imported] - Listeria monocytogenes
 C:Species: Listeria monocytogenes
 C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 C:Accession: AC1382
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Krefte, J.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Me
 ok, C.; Schluster, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
 A.; Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AC1382
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-336 <GLA>
 A:Cross-references: GB:AL592022; PIDN:CA00537.1; PID:g16411947; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: gap
 C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 73.6%; Score 1262; DB 2; Length 336;
 Best Local Similarity 73.3%; Pred. No. 1.1e-83;
 Matches 247; Conservative 33; Mismatches 55; Indels 2; Gaps 2;

QY 1 MVYKVGINGFGRIGLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTEV 60
 Db 1 MTVKVGINGFGRIGLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTEV 60
 QY 61 KEGGFVNGNFTKVSARDPENIDWATDGVETVLEATGFFAKKAAEKHLHANGAKKVI 120
 Db 61 HGFVNGKVKVLANRNPPELPGDGLGVILECTGFFAQDKAEHLHKA-GAKKVI 119
 QY 121 TAPGNDVKTIVFNTNHDILDTETVISGASCTTNCCLAPMAKALHDAFGIOKGLMTTHA 180
 Db 120 SAPATGDMKTIYVNVNHTLDGTETVISGASCTTNCCLAPMAKVLDEKFGVWGLMTTHA 179
 QY 181 YTCQMLDGPGRGDLRRAGAANIVPNSTGAAGAIGLVIPELNGKLDGAARVPVPT 240
 Db 180 YTGQNTLDAPKPGDFRRARAANIIPNTTGAAKAIGEVLPKLGKLDGAARVPVPT 239
 QY 241 GSVELVTLDKXVSDVDEINAAKASN-DSFGYTEPDIYSSDIVGVSGSLFDATQTKV 299
 Db 240 GSLTELVTLDKXVSDVDEINAAKASN-DSFGYTEPDIYSSDIVGVSGSLFDATQTKV 299
 QY 300 MEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
 Db 300 LTVGQDLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 9
 AD1751
 Glyceraldehyde 3-phosphate dehydrogenase homolog gap [imported] - Listeria innocua (stra
 C:Species: Listeria innocua
 C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AD1751
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Krefte, J.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Me
 ok, C.; Schluster, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
 A.; Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AD1751
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-336 <GLA>
 A:Cross-references: GB:AL592022; PIDN:CA097780.1; PID:g16415075; GSPDB:GN00178
 A:Experimental source: strain Clp11262
 C:Genetics:
 A:Gene: gap
 C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 73.5%; Score 1261; DB 2; Length 336;
 Best Local Similarity 73.3%; Pred. No. 1.3e-83;
 Matches 247; Conservative 32; Mismatches 56; Indels 2; Gaps 2;

QY 1 MVYKVGINGFGRIGLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTEV 60
 Db 1 MTVKVGINGFGRIGLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTEV 60
 QY 61 KEGGFVNGNFTKVSARDPENIDWATDGVETVLEATGFFAKKAAEKHLHANGAKKVI 120
 Db 61 HGFVNGKVKVLANRNPPELPGDGLGVILECTGFFAQDKAEHLHKA-GAKKVI 119
 QY 121 TAPGNDVKTIVFNTNHDILDTETVISGASCTTNCCLAPMAKALHDAFGIOKGLMTTHA 180
 Db 120 SAPATGDMKTIYVNVNHTLDGTETVISGASCTTNCCLAPMAKVLDEKFGVWGLMTTHA 179
 QY 181 YTCQMLDGPGRGDLRRAGAANIVPNSTGAAGAIGLVIPELNGKLDGAARVPVPT 240
 Db 180 YTGQNTLDAPKPGDFRRARAANIIPNTTGAAKAIGEVLPKLGKLDGAARVPVPT 239
 QY 241 GSVELVTLDKXVSDVDEINAAKASN-DSFGYTEPDIYSSDIVGVSGSLFDATQTKV 299
 Db 240 GSLTELVTLDKXVSDVDEINAAKASN-DSFGYTEPDIYSSDIVGVSGSLFDATQTKV 299
 QY 300 MEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
 Db 300 LTVGQDLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 10
 C96987
 Glyceraldehyde 3-phosphate dehydrogenase, gene gapC [imported] - Clostridium acetobutyli
 C:Species: Clostridium acetobutylicum
 C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
 C:Accession: C96987
 R:Nolling, J.; Bzeton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
 J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
 A:Reference number: A96900; MUID:21359325; PMID:21359325
 A:Accession: C96987
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-334 <KUR>
 A:Cross-references: GB:AE001437; PIDN:AAK78686.1; PID:g15023589; GSPDB:GN00168
 A:Experimental source: Clostridium acetobutylicum ATCC824
 C:Genetics:
 A:Gene: CAC0709
 C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 73.4%; Score 1259.5; DB 2; Length 334;
 Best Local Similarity 71.9%; Pred. No. 1.6e-83;
 Matches 241; Conservative 40; Mismatches 53; Indels 1; Gaps 1;

C;Genetics:
A;Gene: gap
C;Superfamily: glycerolaldehyde-3-phosphate dehydrogenase

Query Match 69.0%; Score 1183; DB 2; Length 336;
Best Local Similarity 68.2%; Pred. No. 5.3e-78;
Matches 230; Conservative 43; Mismatches 62; Indels 2; Gaps 2;

QY 1 MVVKVINGPGRIGRLAPRRIONVEGVETRLNDLPNMLAHLKYDTTQGRFDGTV 60
DB 1 MAVKVAINGPGRIGRLAPRRIOVEGLEVAVNDLTDMLAHLKYDTTQGRFTGEV 60

QY 61 KEGFVNGNFIKVSABEDPENIDWTDGVEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
DB 61 VDGFRVNGKEVSEFSDASKLPWKDLNIDVLECTGFTYDKDKQAQAHTEA-GAKVLI 119

QY 121 TAPGNDVKTVFNTNHDILDTETVSGACTTNCCLAPMAKALHDAFGIOKGLMTTHA 180
DB 120 SAPATGDLKTVFNTNHOELDGSSETVSGACTTNSLAPVAKVNLNDDFGIVEGLMTTHA 179

QY 181 YTGQMLDGHPRGDLRRARAGANTVPNSTGAAGAIGLVIPELNGKLDGAQRVPEVT 240
DB 180 YTGQNTQDAPHRGDKRRARAENIIPNSTGAAGAIGLVIPEIDKLGGRQVPEVT 239

QY 241 GSVTELVTLDK-NVSVDEINAAKASNDGFTGTEPIVSSDIVGVSGSLFDATQKV 299
DB 240 GSLTELVTLVLEKQDVTVEQVNEAMKASNESFGYTEDEIVSSDVGVGTYGSLFDATQV 299

QY 300 MEVDGSQLVKVSVYDNEMSYTAOLVRLTYFAKIAK 336
DB 300 MSVGRQLVKVAAYDNEMSYTAOLVRLTYLAELSK 336

RESULT 14
T09633
glycerolaldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Lactobacillus
C;Species: Lactobacillus delbrueckii
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C;Accession: T09633
R;Branny, P.; Delatorre, F.; Garel, J.R.
Microbiology 144, 905-914, 1998
A;Title: An operon coding for three glycolytic enzymes in Lactobacillus delbrueckii subsp.
A;Reference number: Z16788; MUID:98240227; PMID:9579064
A;Accession: T09633
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-338

A;Cross-references: EMBL:AJ000339; NID:g2624189; PIDN:CAA04014.1; PID:g2624191
A;Experimental source: subsp. bulgaricus, strain B107
C;Genetics:
A;Gene: gap
C;Superfamily: glycerolaldehyde-3-phosphate dehydrogenase
C;Keywords: gluconeogenesis; glycolysis; oxidoreductase

Query Match 58.5%; Score 1003; DB 2; Length 338;
Best Local Similarity 58.7%; Pred. No. 5.2e-65;
Matches 199; Conservative 45; Mismatches 89; Indels 6; Gaps 3;

QY 1 MVVKVINGPGRIGRLAPRRIONVEGVETRLNDLPNMLAHLKYDTTQGRFDG 56
DB 1 MTVKVINGPGRIGRLAPRRIMDLGEETKDEIVVAINDLTPAHLAHLKYDSTHGFTH 60

QY 57 TVEVKEGFEVNGNFIKVSABEDPENIDWTDGVEIVLEATGFFAKKEAAEKHLHANGA 115
DB 61 EVSATEDSLVDGKKRYVYAFQPAQNTPEWVNDGVDFVLECTGFTYTSKAKESQAHLDA-GA 119

QY 116 KKVITAPGNDVKTVFNTNHDILDTETVSGACTTNCCLAPMAKALHDAFGIOKGLM 175
DB 120 KKVILAPAGNDLKTIVSVNODTLTADDTIVSAGCTTNSLAPWANALNKEFGIQVGTM 179

QY 176 TTIHAYTATQKVLDPGRGNFRNARAAAENIIPNSTGAAGAIGLVIPELNGKLDGAQR 235

DB 180 TTIHAYTATQKVLDPGRGNFRNARAAAENIIPNSTGAAGAIGLVIPELNGKLDGAQR 239

QY 236 VVPVPTGSEVTLVTLDKNVSVDEINAAKASNDGFTGTEPIVSSDIVGVSGSLFPAT 295

DB 240 VPKDGSSETLVTILDKKVTAEVNAAMKYESPSFAYNADQIVSTDVGLMTAGSIFDPT 299

QY 296 QTKWNEVDGSQLVKVSVYDNEMSYTAOLVRLTYFAKIAK 334

DB 300 QIQVITAGDKQLVKVYDNEYSTFCQVRLTLHFAFL 338

RESULT 15
F90881
glycerolaldehyde 3-phosphate dehydrogenase C [imported] - Escherichia coli (strain O157:H7)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: F90881
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gatawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinsagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90881
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-333

A;Cross-references: GB:BA000007; PIDN:BA035445.1; PID:g13361488; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECS2022
C;Superfamily: glycerolaldehyde-3-phosphate dehydrogenase

Query Match 57.9%; Score 993; DB 2; Length 333;
Best Local Similarity 59.3%; Pred. No. 2.7e-64;
Matches 198; Conservative 54; Mismatches 76; Indels 6; Gaps 5;

QY 4 KYGNGFGRIGRLAPRRIONVE-GVEVTRINDLTPNMLAHLKYDTTQGRFDGTVVEKE 62

DB 3 KYGNGFGRIGRLVLRLLVKNIDVVAINDLTPKILAYLLKHDNSYGVGFPWSDVDFTE 62

QY 63 GGFVNGNFIKVSABEDPENIDWTDGVEIVLEATGFFAKKEAAEKHLHANGAKKVI 122

DB 63 DSLIVDGRKSIAYAEKAKNIPWAKAGAEIIVECTGFVTSAEKSAHLDA-GAKKVLISA 121

QY 123 PGNDVKTVFNTNHDILDTETVSGACTTNCCLAPMAKALHDAFGIOKGLMTTHAYT 182

DB 122 PAG-EMKTVYKVNDDTLGNDTIVSVASCTTNCCLAPMAKALHDSFGIEVGTMTTHAYT 180

QY 183 GQOMILDGPHRGDLRRARAGANTVPNSTGAAGAIGLVIPELNGKLDGAQAQRVVPVTCG 242

DB 181 GTQSLVDGP-RGKDLRASAFAAENIIPHTTGAAGAIGLVIPELSGLKAGHAQRVVPVTCG 239

QY 243 VTELVTLDKNVSVDEINAAKAA--SNDSPGYTDEPIVSSDIVGVSGSLFDATQTKVM 300

DB 240 VTELVSILGKVTAEVNNALQAATTNNESTGYTDEEIVSSDIIGSHFGSVFDTATQTEIT 299

QY 301 EVDGSQLVKVSVYDNEMSYTAOLVRLTYFAKIAK 334

DB 300 AVGDIQLVKTVYDNEYGFVTLRTLEKFAKL 333

Search completed: May 11, 2004, 16:49:49
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 16:42:06 ; Search time 18 Seconds
(without alignments)
971.976 Million cell updates/sec

Title: US-10-650-369-12

Perfect score: 1715

Sequence: 1 MVVKVINGFRIGRLAFRR.....EMSYTAQLVRLTYEYFAKIAK 336

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1709	99.7	335	1	G3P_STRPY P50467 streptococ
2	1705	99.4	335	1	G3P_STRP3 Q8k8m9 streptococ
3	1612	94.0	335	1	G3P_STREQ Q59906 streptococ
4	1338.5	78.0	337	1	G3P_LACLA P52987 lactococcc
5	1289.5	75.2	334	1	G3P_CLOPA Q59309 clostridium
6	1259.5	73.4	334	1	G3P_CLOAB Q52631 clostridium
7	1190	69.4	336	1	G3P1_STARP Q8cpv5 staphylococ
8	1183	69.0	336	1	G3P1_STAAE Q32755 lactobacill
9	1003	58.5	338	1	G3P_LACDE Q32755 lactobacill
10	993	57.9	333	1	G3P3_EC057 P58072 escherichia
11	921	57.8	333	1	G3P3_EC011 P33898 escherichia
12	921	53.7	337	1	G3P_MYCPN P75358 mycoplasma
13	908.5	53.0	334	1	G3P_CORGL Q01651 corynebacte
14	908	52.9	337	1	G3P_MYCGE P47543 mycoplasma
15	891.5	52.0	334	1	G3P1_BACSU P09124 bacillus su
16	886.5	51.7	334	1	G3P_BACME P23722 bacillus me
17	880.5	51.3	334	1	G3P_BACST P00362 bacillus st
18	873	50.9	332	1	G3P_THEMA P17721 thermotoga
19	863	50.3	335	1	G3P_BORBU P46795 borrelia bu
20	843	49.2	342	1	G3P_AQUAE O87161 aquifex aeo
21	842.5	49.1	336	1	G3P_STRCO Q2518 streptomyc
22	817	47.6	330	1	G3P1_SALTY P24165 salmonella
23	816	47.6	339	1	G3P_MYCTU O06822 mycobacteri
24	815	47.5	339	1	G3P_MYCLE P46713 mycobacteri
25	813	47.4	330	1	G3P1_EC011 P06977 escherichia
26	807.5	47.1	339	1	G3P_EAELN P44304 haemophilus
27	801	46.7	339	1	G3P_MYCAV P94915 mycobacteri
28	794.5	46.3	337	1	G3P1_RHIRA Q9c136 rhizomucor
29	794	46.3	330	1	G3PC_TRYBB P10097 trypanosoma
30	787.5	45.9	332	1	G3P1_SCHPO P78958 schizosacch
31	784.5	45.7	336	1	G3P1_RALSO P52694 ralstonia s
32	783.5	45.7	337	1	G3P_MONAN P53430 monascus an
33	779.5	45.5	331	1	G3P_THEAO P00361 thermus aqu

34	779.5	45.5	337	1	G3P_Clapu	Q00584 claviceps p
35	777.5	45.3	332	1	G3P_STRAU	Q59800 streptomyc
36	775	45.2	336	1	G3P2_SYNY3	P85005 synchocyst
37	772	45.0	333	1	G3P_STRAE	P54226 streptomyc
38	772	45.0	337	1	G3P2_ANASP	P58554 anabaena sp
39	771.5	45.0	337	1	G3P_COCHE	P29497 cochllobolu
40	771	45.0	330	1	G3PC_LEIME	Q01558 leishmania
41	770.5	44.9	335	1	G3P2_SCHPO	Q43026 schizosacch
42	770.5	44.9	337	1	G3P_CURLU	P28844 curvularia
43	768.5	44.8	336	1	G3P_ASPNG	Q12552 aspergillus
44	766.5	44.7	336	1	G3P_EMENI	P20445 emericeella
45	762.5	44.5	338	1	G3P_NEUCR	P54118 neurospora

ALIGNMENTS

RESULT 1

ID	G3P_STRPY	STANDARD;	PRT;	335 AA.
AC	P50467;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)			
DE	(Plasminogen-binding protein) (Plasmin receptor).			
GN	GAP OR FLR OR GAPA OR SP10274 OR SPYMI8_0261.			
OS	Streptococcus pyogenes, and			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=1314, 186103;			
RN	[1]_TaxID=1314, 186103;			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-73; 160-173 AND 186-216.			
RC	STRAIN=64/14;			
RX	MEDLINE=92355491; PubMed=1322883;			
RA	Lottenberg R., Broder C.C., Boyle M.D., Kain S.J., Schroeder B.L.,			
RA	Curtiss R. III;			
RT	"Cloning, sequence analysis, and expression in Escherichia coli of a			
RT	streptococcal plasmin receptor.";			
RL	J. Bacteriol. 174:5204-5210 (1992).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN=SF370 / ATCC 700294 / Serotype M1;			
RC	MEDLINE=21192684; PubMed=11296296;			
RX	Ferratti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,			
RA	Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,			
RA	Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,			
RA	Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;			
RT	"Complete genome sequence of an M1 strain of Streptococcus			
RT	pyogenes.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663 (2001).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MGAS6232 / Serotype M18;			
RX	MEDLINE=21927593; PubMed=11917108;			
RA	Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,			
RA	Sylvia G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,			
RA	Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,			
RA	Kapur V., Daly J.A., Veasy L.G., Musser J.M.;			
RT	"Genome sequence and comparative microarray analysis of serotype M18			
RT	group A Streptococcus strains associated with acute rheumatic fever			
RT	outbreaks.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).			
RN	[4]			
RP	SEQUENCE OF 1-11; 20-30; 103-128; 162-171 AND 199-215.			
RC	STRAIN=JRS4 / Serotype M6;			
RA	Hogan D.A., Du P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,			
RA	VanBogelen R.A.;			
RT	"Two-dimensional gel electrophoresis map of Streptococcus pyogenes			
RT	proteins.";			
RL	Submitted (MAY-2000) to Swiss-Prot.			
CC	-!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN.			

CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M95569; AAA26953.1; --
CC EMBL; AE006434; AAK33348.1; --
CC EMBL; AE009973; AAL97041.1; --
CC DR HSSP; P00362; 1GD1.
CC DR InterPro; IPR001173; GAP_dhhydrogenase.
CC DR InterPro; IPR006424; GAPDH-I.
CC DR Pfam; PF00044; gpdh; 1.
CC DR Pfam; PF02800; gpdh; C; 1.
CC DR PRINTS; PR00078; G3PDHGRGNASE.
CC DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC DR PROSITE; PS00071; GAPDH; 1.
CC KW Glycolysis; Oxidoreductase; NAD; Complete proteome.
CC FT INIT MET 0
CC FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE (BY
CC SIMILARITY)
CC FT ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS
CC (BY SIMILARITY).
CC FT CONFLICT 261 261 A -> S (IN REF. 1).
CC FT SEQUENCE 335 AA; 35811 MW; F06006EE253C8A3F CRC64;
CC
CC Query Match 99.7%; Score 1709; DB 1; Length 335;
CC Best Local Similarity 99.7%; Pred. No. 3.4e-111; Indels 0; Gaps 0;
CC Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 2 VVKVINGFGFRIGRLAFRIQVGEVTRINDLTPNMLAHLKYDTTQGRFDGTVGVK 61
CC Db 1 VVKVINGFGFRIGRLAFRIQVGEVTRINDLTPNMLAHLKYDTTQGRFDGTVGVK 60
CC QY 62 EGGFEVNGFNFKVSAERDPENIDWATDGVIEATGFFAKKEAAEKHLHANGAKKVIT 121
CC Db 61 EGGFEVNGFNFKVSAERDPENIDWATDGVIEATGFFAKKEAAEKHLHANGAKKVIT 120
CC QY 122 APGGNDVKTIVFNTNHDILDTGVISGASCTTNCLAPMAKALHDAFGIQKGLMTTHAY 181
CC Db 121 APGGNDVKTIVFNTNHDILDTGVISGASCTTNCLAPMAKALHDAFGIQKGLMTTHAY 180
CC QY 182 TGDQMLDGPGRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAARQVPVPTG 241
CC Db 181 TGDQMLDGPGRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAARQVPVPTG 240
CC QY 242 SVTELVTLLDKNVSVDENAAKASNDSPGYTEDPIVSSDIVGSVGSIFDQTQKME 301
CC Db 241 SVTELVTLLDKNVSVDENAAKASNDSPGYTEDPIVSSDIVGSVGSIFDQTQKME 300
CC QY 302 VDGSQLVKVSNYDNEMSYTQCLVRLTFEYFAKIAK 336
CC Db 301 VDGSQLVKVSNYDNEMSYTQCLVRLTFEYFAKIAK 335

RESULT 2

G3P_STRP3 STANDARD; PRT; 335 AA.
ID G3P_STRP3
AC Q8K8M9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)

DE (Plasminogen-binding protein) (Plasmin receptor).
GN GAP OR PLR OR SPYK3_0201 OR SPS0207.
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAINMGRS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G.B., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Perkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Mueser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SSI-1 / Serotype M3;
RX MEDLINE=22683278; PubMed=12799345;
RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
RA Okaguchi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA Hayashi H., Hattori M., Hamada S.;
RT "Genome-scale sequencing of an M3 strain of Streptococcus reveals a
RT large-scale genomic rearrangement in invasive strains and new insights
RT into phage evolution";
RL Genome Res. 13:1042-1055(2003).
CC -!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN (By similarity).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE014140; AM78808.1; --
CC EMBL; AP005141; BAC63302.1; --
CC InterPro; IPR000173; GAP_dhhydrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh; C; 1.
CC PRINTS; PR00078; G3PDHGRGNASE.
CC DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC DR PROSITE; PS00071; GAPDH; 1.
CC KW Glycolysis; Oxidoreductase; NAD; Complete proteome.
CC FT INIT MET 0
CC FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE (BY
CC SIMILARITY)
CC FT ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS
CC (BY SIMILARITY).
CC FT SEQUENCE 335 AA; 35841 MW; 4DCB76F382F6F698 CRC64;
CC
CC Query Match 99.4%; Score 1705; DB 1; Length 335;
CC Best Local Similarity 99.4%; Pred. No. 6.3e-111;
CC Matches 333; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
CC
CC QY 2 VVKVINGFGFRIGRLAFRIQVGEVTRINDLTPNMLAHLKYDTTQGRFDGTVGVK 61
CC Db 1 VVKVINGFGFRIGRLAFRIQVGEVTRINDLTPNMLAHLKYDTTQGRFDGTVGVK 60
CC QY 62 EGGFEVNGFNFKVSAERDPENIDWATDGVIEATGFFAKKEAAEKHLHANGAKKVIT 121

```

Db 61 EGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHTNGAKKVIT 120
QY 122 APGNDVKTVVFNTHDILDTGTEVTSAGSCTTNCLAPMAKALHDAFGIQKGLMTTIHAY 181
Db 121 APGNDVKTVVFNTHDILDTGTEVTSAGSCTTNCLAPMAKALHDAFGIQKGLMTTIHAY 180
QY 182 TGDQMLDGHRRGDLRRRAGAAANIVPNSGTGAAGAIGLVIPELNGKLDGAARVPVPTG 241
Db 181 TGDQMLDGHRRGDLRRRAGAAANIVPNSGTGAAGAIGLVIPELNGKLDGAARVPVPTG 240
QY 242 SVTELVTLDKNVSVDEINAAKASNDSPGYTEDPIVSSDIVGVSYGSLFDATQTKWME 301
Db 241 SVTELVTLDKNVSVDEINAAKASNDSPGYTEDPIVSSDIVGVSYGSLFDATQTKWME 300
QY 302 VDSQSLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 VDSQSLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 335

RESULT 3
G3P_STREQ STANDARD; PRT; 335 AA.
AC Q59906;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)
DE (Plasminogen-binding protein) (Plasmin receptor).
GN GAP OR GAPC.
OS Streptococcus equisimilis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=119602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H46A;
RA MEDLINE=96305364; PubMed=8706717;
RX Gase K., Gase A., Schirmer H., Malke H.;
RT "Cloning, sequencing and functional overexpression of the
RT Streptococcus equisimilis H46A gapC gene encoding a
RT glyceraldehyde-3-phosphate dehydrogenase that also functions as a
RT plasmin(ogen)-binding protein. Purification and biochemical
RT characterization of the protein."
RL Eur. J. Biochem. 239:42-51(1996).
CC -|- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN.
CC -|- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -|- PATHWAY: Second phase of glycolysis; first step.
CC -|- SUBUNIT: Homotetramer (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; X97788; CAA66377.1; -
EMBL; Y12602; CAA73174.1; -
PIR; S71350; S71350.
HSP; P00362; IGD1.
InterPro; IPR000173; GAP_dhhydrogenase.
InterPro; IPR006424; GAPDH-1.
Pfam; PF00044; gpdh; 1.
Pfam; PF02800; gpdh; C1.
PRINTS; PR00078; G3PDHGRGNASE.
TRGFAMs; TIGR01534; GAPDH-1; 1.
PROSITE; PS00071; GAPDH; 1.
Glycolysis; Oxidoreductase; NAD.

```

```

FT INIT MET 0 0 BY SIMILARITY.
FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE (BY
FT SIMILARITY).
FT ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS
FT (BY SIMILARITY).
SQ SEQUENCE 335 AA; 35739 MW; PE7ACFDPE7663E46 CRC64;

Query Match 94.0%; Score 1612; DB 1; Length 335;
Best Local Similarity 95.5%; Pred. No. 1.7e-104; Indels 0; Gaps 0;
Matches 320; Conservative 2; Mismatches 13;

QY 2 VVKVINGFGIGRIGLAFRRIONVEGVETRIINDLTDNMLLAHLKLYDTTQGRFGDTVEVK 61
Db 1 VVKVINGFGIGRIGLAFRRIONVEGVETRIINDLTDNMLLAHLKLYDTTQGRFGDTVEVK 60
QY 62 EGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVIT 121
Db 61 EGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKPLHANGAKKVIT 120
QY 122 APGNDVKTVVFNTHDILDTGTEVTSAGSCTTNCLAPMAKALHDAFGIQKGLMTTIHAY 181
Db 121 APGNDVKTVVFNTHDILDTGTEVTSAGSCTTNCLAPMAKALHDAFGIQKGLMTTIHAY 180
QY 182 TGDQMLDGHRRGDLRRRAGAAANIVPNSGTGAAGAIGLVIPELNGKLDGAARVPVPTG 241
Db 181 TGDQMLDGHRRGDLRRRAGAAANIVPNSGTGAAGAIGLVIPELNGKLDGAARVPVPTG 240
QY 242 SVTELVTLDKNVSVDEINAAKASNDSPGYTEDPIVSSDIVGVSYGSLFDATQTKWME 301
Db 241 SVTELVTLDKNVSVDEINAAKASNDSPGYTEDPIVSSDIVGVSYGSLFDATQTKWME 300
QY 302 VDSQSLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 VDSQSLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 335

RESULT 4
G3P_LACLA STANDARD; PRT; 337 AA.
ID P52987;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAP OR LD0559.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LM230;
RX MEDLINE=95291425; PubMed=7773380;
RA Cancilla M.R., Hillier A.J., Davidson B.E.;
RT "Lactococcus lactis glyceraldehyde-3-phosphate dehydrogenase gene,
RT gap: further evidence for strongly biased codon usage in glycolytic
RT pathway genes."
RL Microbiology 141:1027-1036(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarre K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis il1403."
RL Genome Res. 11:731-753(2001).
CC -|- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -|- PATHWAY: Second phase of glycolysis; first step.
CC -|- SUBUNIT: Homotetramer (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; L36907; AAC41453.1; -.
CC EMBL; AE006290; AAK04657.1; -.
CC F1R; G86694; G86694.
CC HSP; P17721; 1HDG.
CC InterPro; IPR000173; GAP dhdrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh; 1.
CC PRINTS; PR00078; G3PDHGRGNASE.
CC TIGRfam; TIGR01534; GAPDH-I; 1.
CC PROSITE; PS00071; GAPDH; Complete proteome.
KW GLYCERALDEHYDE 3-PHOSPHATE
FT BINDING 152 152
FT ACT_SITE 179 179
FT CONFLICT 143 143
FT SEQUENCE 337 AA; 36057 MW; 175B8C6AAEF5F59D CRC64;
Query Match 78.0%; Score 1338.5; DB 1; Length 337;
Best Local Similarity 77.2%; Pred. No. 1.4e-85;
Matches 260; Conservative 31; Mismatches 45; Indels 1; Gaps 1;
QY 1 MVVKGINGFGRIQLAPRRQNVGEVVEVTRINDLTPNMLAHLKYDTTQGRFGDTVEV 60
DB 1 MVVKGINGFGRIQLALRRQVEGEVEVAHINDLTPNMLAHLKYDTTQGRFGDTVEV 60
QY 61 KEGFVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKAAEKLHANGAKKVV 120
DB 61 KEGFVNGKVKVTAERNPEDIQWADSGVEIVLEATGFFATKEAEKHLHPCGAKVLI 120
QY 121 TAPGNDVKTIVFNTNHDILDTGTTVSGASCTTNCPLAPMAKALHDAFGIQGLMTTII 180
DB 121 TAPGNDVKTIVFNTNHDILDTGTTVSGASCTTNCPLAPMAKALHDAFGIQGLMTTII 180
QY 181 YTGQMLDGPGRGDLRRARAGAAIVPNSGTGAAGAIGLVIPELNGKLDGAQAQRPVPT 240
DB 181 YTGQMLDGPGRGDLRRARAAENIVPASSGAAGAIGLVIPELNGKLDGAQAQRPVPT 240
QY 241 GSVTELVTLTKNVSVDEINAAKAAANDSGFYTEDPIVSSDIVGVSGSLFDATQTKVM 300
DB 241 GSVTELVTLTKNVSVDEINAAKAAANDSGFYTEDPIVSSDIVGVSGSLFDATQTKVM 300
QY 301 EV-DGSQLVKVSVYDNEMSVTAQLVRLTLEYFAKIAK 336
DB 301 DLKGGQLVKAAYDNEMSVTAQLVRLTLEYFAKIAK 337
RESULT 5
G3P_CLOPA
ID G3P_CLOPA STANDARD; PRT; 334 AA.
AC Q59309;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH) (CP
DE 17/CP 18).
GN GAP.
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1501;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RA Oster T., Assobhei O., Scherrer S., Brantant G., Brantant C.;
RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP PARTIAL SEQUENCE OF 1-26.
RC STRAIN=W5;
RX MEDLINE=98291870; PubMed=9629918;
RA Flengrud R., Skjeldal L.;
RT "Two-dimensional gel electrophoresis separation and N-terminal
RT sequence analysis of proteins from Clostridium pasteurianum W5.";
RL Electrophoresis 19:802-806(1998).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (by similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; X72219; CAA51020.1; -.
CC F1R; S34254; S34254.
CC HSP; P00362; 1GD1.
CC InterPro; IPR000173; GAP dhdrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh; 1.
CC PRINTS; PR00078; G3PDHGRGNASE.
CC TIGRfam; TIGR01534; GAPDH-I; 1.
CC PROSITE; PS00071; GAPDH; 1.
KW GLYCOLYSIS; Oxidoreductase; NAD.
FT BINDING 150 150
FT ACT_SITE 177 177
FT SEQUENCE 334 AA; 36078 MW; D15905D0DA7F62E7 CRC64;
Query Match 75.2%; Score 1289.5; DB 1; Length 334;
Best Local Similarity 74.3%; Pred. No. 3.4e-82;
Matches 249; Conservative 33; Mismatches 52; Indels 1; Gaps 1;
QY 2 VVKVGINGFGRIQLAPRRQNVGEVVEVTRINDLTPNMLAHLKYDTTQGRFGDTVEV 61
DB 1 MTKVAINGFGRIQLALRRILEVPLEVVAINDLTPNMLAHLKYDSSQGRFGEIEV 60
QY 62 EGGFVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKAAEKLHANGAKKVVIT 121
DB 61 EGAFVNGKVKVFAEADPEKLPNGELGIDVLECTGFTKKKAAEAHVRA-GAKKVVIS 119
QY 122 AFGNDVKTIVFNTNHDILDTGTTVSGASCTTNCPLAPMAKALHDAFGIQGLMTTII 181
DB 122 AFGNDLKTIVFNTNHDILDTGTTVSGASCTTNCPLAPMAKALHDAFGIQGLMTTII 179
QY 182 TGDQMLDGPGRGDLRRARAGAAIVPNSGTGAAGAIGLVIPELNGKLDGAQAQRPVPT 241
DB 180 TNDQNTLDGPGRGDLRRARAAVSIIPNSTGAAGAIAQVPELNGKLDGAQAQRPVPT 239
QY 242 SVTELVTLTKNVSVDEINAAKAAANDSGFYTEDPIVSSDIVGVSGSLFDATQTKME 301
DB 240 SVTELVTLTKNVSVDEINAAKAAANDSGFYTEDPIVSSDIVGVSGSLFDATQTKME 299
QY 302 VDSQLVKVSVYDNEMSVTAQLVRLTLEYFAKIAK 336
DB 300 VDSQLVKVSVYDNEMSVTAQLVRLTLEYFAKIAK 334
RESULT 6
G3P_CLOAB
ID G3P_CLOAB STANDARD; PRT; 334 AA.

```

```

052631;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAP OR GAPC OR CAC0709.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
[1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=ATCC 824 / DSM 792 / VMM B-1787;
RX MEDLINE=9339246; PubMed=10463150;
RA Schreiber W., Durre P.;
RT "The glyceraldehyde-3-phosphate dehydrogenase of Clostridium
acetobutylicum: Isolation and purification of the enzyme, and
sequencing and localization of the gap gene within a cluster of other
glycolytic genes.";
RL Microbiology 145:1839-1847(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VMM B-1787;
RX MEDLINE=9339246; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabache F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate +
NAD(+) = 3-phospho-D-glyceroyl phosphate + phosphate +
NADH.
CC -!- PATHWAY: Second phase of glycolysis, first step.
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
dehydrogenase family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to licenses@isb-sib.ch).
-----
EMBL; AF043386; AAC13160.1; -.
ENWL; AE007586; AAK78686.1; -.
PR; C96987; C96987.
HSP; P17721; IHG.
InterPro; IPR000173; GAP_dhdnrgenase.
InterPro; IPR006424; GAPDH-I.
Pfam; PF00044; gpdh; 1.
Pfam; PF02800; gpdh.C; 1.
PRINTS; PR00078; GFPDHDGRNAS.
TIGRams; TIGR01534; GAPDH-I; 1.
PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductases; NAD; Complete proteome.
FT BINDING 150 177 GLYCERALDEHYDE 3-PHOSPHATE.
FT ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ SEQUENCE 334 AA; 35850 MW; 10C52A174BE789B5 CRC64;
Query Match 73.4%; Score 1259.5; DB 1; Length 334;
Best Local Similarity 71.9%; Pred. No. 4e-80;
Matches 241; Conservative 40; Mismatches 53; Indels 1; Gaps 1;
QY 2 VKVKGNGFGRIQLAFRRIONVEGVETVRINDLTDPNMLAHLLKYDTTCGRFDGTVEVK 61
Db 1 MAKIANGFRIQLALRRILEVPGLVEVAINDLTDKMLAHLFKVDSSGGFRNGIEVK 60
QY 62 EGGEFVNNGPIKVSARDPNIDWATGDVIEVLTEATGFPAKEAAEKHLHANGAKKVIT 121

```

```

Best Local Similarity 69.4%; Pred. No. 2.6e-75;
Matches 234; Conservative 39; Mismatches 62; Indels 2; Gaps 2;

QY 1 MVVKGINGFGRIGLAFRRQNVGVEVTRINDLTPNMLAHLKYDTTQGRPDGTVEV 60
Db 1 NAIKVAINGFGRIGLAFRRQNVGVEVTRINDLTPNMLAHLKYDTTQGRPDGTVEV 60
QY 61 KEGFVNGNFIKVAERDPENIDWATDGVIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 LEGFVNGNFIKVAERDPENIDWATDGVIVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVFNTNHDILDTETVIGSACTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180
Db 120 SAPAKGDKVTIVFNTNHDILDTETVIGSACTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180
QY 181 YTGDMILDGPHRGDGLRRARAGANIVPNSGTGAAGAIGLVIPELNGKLDGAAQORVPVPT 240
Db 180 YTGDMILDGPHRGDGLRRARAGANIVPNSGTGAAGAIGLVIPELNGKLDGAAQORVPVPT 240
QY 241 GSVTELVTLTK-NVSVDEINAAKAAANDSGFTGTEDEIVSSDIVGMYGSLFDATQTKV 299
Db 240 GSVTELVTLTK-NVSVDEINAAKAAANDSGFTGTEDEIVSSDIVGMYGSLFDATQTKV 299
QY 300 MEVDGSQLVKVSWYDNEVMSYTAQLVRLTYLFAKIAK 336
Db 300 MTVGDRQLVKVAAWYDNEVMSYTAQLVRLTYLFAKIAK 336

RESULT 8
ID G3P1 STAAAM STANDARD; PRT; 336 AA.
AC Q925C5;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE Glyceraldehyde 3-phosphate dehydrogenase 1 (EC 1.2.1.12) (GAPDH 1).
GN GAPA OR GAP OR SAV0772 OR SA0727 OR MW0734.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315),
OS Staphylococcus aureus (strain MW2), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 196620, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BB;
RA Morrissey J.A., Williams P.;
RT "Isolation and characterisation of a glycolytic operon in
RT Staphylococcus aureus.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699, and N315;
EX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
EX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwano N., Asano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";

```

```

RL Lancet 359:1819-1827 (2002).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AJ133520; CAB38645.1; -
CC EMBL; AF003360; BAB5834.1; -
CC EMBL; AF003131; BAB41960.1; -
CC EMBL; AF004824; BAB94599.1; -
CC FIR; E89850; E89850.
CC SWISS-2DPAGE; Q925C5; STAAAM.
CC HSSP; P17721; LHG.
CC InterPro; IPR006424; GAPDH-I.
CC InterPro; IPR000173; GAP_dhydrogenase.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh; 1.
CC PRINTS; PR00078; G3PDHGRNASE.
CC TIGRFS; TIGR01534; GAPDH-1; 1.
CC PROSITE; PS00071; GAPDH; 1.
CC Glycolysis; Oxidoreductase; NAD; Complete proteome
CC BINDING 151 151
CC SIMILARITY.
CC ACT_SITE 178 178
CC ACTIVATES THIOL GROUP DURING CATALYSIS
CC (BY SIMILARITY).
CC
CC SEQUENCE 336 AA; 36281 MW; 37A6CEA9376779E5 CRC64;
CC
CC Query Match 69.0%; Score 1183; DB 1; Length 336;
CC Best Local Similarity 68.2%; Pred. No. 7.9e-75;
CC Matches 230; Conservative 43; Mismatches 62; Indels 2; Gaps 2;
QY 1 MVVKGINGFGRIGLAFRRQNVGVEVTRINDLTPNMLAHLKYDTTQGRPDGTVEV 60
Db 1 NAIKVAINGFGRIGLAFRRQNVGVEVTRINDLTPNMLAHLKYDTTQGRPDGTVEV 60
QY 61 KEGFVNGNFIKVAERDPENIDWATDGVIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 LEGFVNGNFIKVAERDPENIDWATDGVIVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVFNTNHDILDTETVIGSACTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180
Db 120 SAPATGDLKTIIVFNTNHDILDTETVIGSACTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180
QY 181 YTGDMILDGPHRGDGLRRARAGANIVPNSGTGAAGAIGLVIPELNGKLDGAAQORVPVPT 240
Db 180 YTGDMILDGPHRGDGLRRARAGANIVPNSGTGAAGAIGLVIPELNGKLDGAAQORVPVPT 240
QY 241 GSVTELVTLTK-NVSVDEINAAKAAANDSGFTGTEDEIVSSDIVGMYGSLFDATQTKV 299
Db 240 GSVTELVTLTK-NVSVDEINAAKAAANDSGFTGTEDEIVSSDIVGMYGSLFDATQTKV 299
QY 300 MEVDGSQLVKVSWYDNEVMSYTAQLVRLTYLFAKIAK 336
Db 300 MTVGDRQLVKVAAWYDNEVMSYTAQLVRLTYLFAKIAK 336

RESULT 9
G3P_LACDE
ID G3P_LACDE STANDARD; PRT; 338 AA.
AC G32755;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

DT	28-FEB-2003 (Rel. 41, Last annotation update)	
DE	Glyceraldehyde 3-phosphate dehydrogenase [SC 1.2.1.12] (GAPDH).	
GN	GAP.	
OS	Lactobacillus delbrueckii (subsp. bulgaricus).	
OC	Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;	
OC	Lactobacillus.	
OX	NCBI_TaxID=1585;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=B107;	
RX	MEDLINE=98240227; PubMed=9579064;	
RA	Branny P., Delatorre F., Garel J.R.;	
RT	"An operon encoding three glycolytic enzymes in Lactobacillus	
RT	delbrueckii subsp. bulgaricus: glyceraldehyde-3-phosphate	
RT	dehydrogenase, phosphoglycerate kinase and triosephosphate	
RT	isomerase.";	
RL	Microbiology 144:905-914 (1998).	
CC	-1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +	
CC	NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.	
CC	-1- PATHWAY: Second phase of glycolysis; first step.	
CC	-1- SUBUNIT: Homotrimer (By similarity).	
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).	
CC	-1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate	
CC	dehydrogenase family.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/	
CC	or send an email to license@sb-sib.ch).	
CC	-----	
DR	EMBL; AJ000339; CAA04014.1; .	
DR	PIR; T09633; T09633.	
DR	HSP; P17721; IHG.	
DR	InterPro; IPR000173; GAP_dhhydrogenase.	
DR	InterPro; IPR006424; GAPDH-1.	
DR	Pfam; PF00044; gpdh; 1.	
DR	Pfam; PF02800; gpdh; C; 1.	
DR	PRINTS; PR00078; G3PDHGRGNASE.	
DR	TIGRfam; TIGR01534; GAPDH-1.1.	
DR	PROSITE; PS00071; GAPDH; FALSE_NEG.	
KW	Glycolysis; NAD; Oxidoreductase.	
FT	GLYCERALDEHYDE 3-PHOSPHATE	
FT	GLYCOLYSIS; NAD; Oxidoreductase.	
FT	ACT_SITE 183 183	
FT	ACTIVATES THIOL GROUP DURING CATALYSIS	
FT	(BY SIMILARITY).	
FT	(BY SIMILARITY).	
SQ	SEQUENCE 338 AA; 36564 MW; DP9758C9CA4F7DFA CRC64;	
	Query Match 58.5%; Score 1003; DB 1; Length 338;	
	Best Local Similarity 58.7%; Pred. No. 2.2e-62;	
	Matches 199; Conservative 45; Mismatches 89; Indels 6; Gaps 3;	
QY	1 MVKVGINGRIGRLAFRI---QNVGEVETRINDLTPNMHLHLYDVTQGRFDG 56	
DB	1 MTWKIGINGRIGRLAFRIIMDLGETKDIEVVAINDLTPMAMHLHLYDTHGTGTFDH 60	
QY	57 TVEYKEGGEFVNGNFIKVSARDPENIDWA-TDGEIVLEATGTFPAKKAEEKHLHANGA 115	
DB	61 EVSATEDSLVVDGKKRYVAEPQNIPIWKNDCGVDFLECTGFTYSKAKSQAHDA-GA 119	
QY	116 KKVITAPGNDVKTVFNTNHDILDGTEIVVISGASCTTNCIAPMAKALHDAFGI QKGLM 175	
DB	120 KRVLISAPAGNDKTVIYVSNQDPTLTADDTIVSAGSCTTISLAPMALNKEFGIQVGTM 179	
QY	176 TTHAYTGDQIMLDGPHRGDLRRARAGAAIVPNSTGAAKAIGLVIPELNGKLDGAQR 235	
DB	180 TTHAYTATQKVLDPGRGNFRNRAAAENIIPSTGAAKAIGLVIPELNGKLDGHAQR 239	
QY	236 VPVPTGVTSELVLTLDKXVSDEINAAKASNDSPGYTEDPIVSSDIVGVSYGSLFDT 295	
DB	240 VPVKDGETELVTLTDKKVTAEEVNAAMKKYESSPFAYNADQIVSTDLVGMTAGSI FDDPT 299	

```
FT ACT_SITE 177 177 ACTIVATES THIOLE GROUP DURING CATALYSIS
FT SEQUENCE 333 AA; 35763 MW; A2F77CB2E773B64C CRC64;
SQ
Query Match 57.9%; Score 993; DB 1; Length 333;
Best Local Similarity 59.3%; Pred. No. 1.le-61;
Matches 198; Conservative 76; Mismatches 6; Gaps 5;
QY 4 KVGINGGRTGLAFRIQNV-EQVEVTRINDLTDPMNLALHILKYDTTQGRPDGTVEVKE 62
DB 3 KVGINGGRTGLAFRIQNV-EQVEVTRINDLTDPMNLALHILKYDTTQGRPDGTVEVKE 62
QY 63 GGFVNGNFKIVSAERDENIDATGVEIVLEATGPFAXKAEAEKHLHANGAKKVVITA 122
DB 63 DSLIVDKSTAVYAEKAKIIPWAKAGAEIVECTGYTSAEKSQAHLD-BAKVLISA 121
QY 123 PGGNVKTIVVNTNHDILDTGTTETVSGASTTNCCLAPAKALHDAFGIQKGLMTTHAYT 182
DB 122 PAG-EMKTIYVKVNDTLDGNDTIVSVASCTTNCCLAPAKALHDSFGIEVGTTHAYT 180
QY 183 GDQMLDGGHGGDLRRARAGANIVNSTGAKAIGLVIPELNGKLDGAAQRPVPTGS 242
DB 181 GTQSLVDGP-RGKDLRASRAAENIIPHTOAKAIGLVIPELNGKLDGAAQRPVPTGS 239
QY 243 VTELVTLDKNVSVDEINAAKAA--SNDGFGYTTEDPIVSSDIVGVSGSLFDATQTKVM 300
DB 240 VTELVSILGKVKTAEEVNNALKQATTNNEGFGYTTEDPIVSSDIVGVSGSLFDATQTEIT 299
QY 301 EYDSQVSVKVSVDNEMSYTAQIVRTLEYFAKI 334
DB 300 AVGLQLVKTVAVYDNEYGFVTQIIRLEKFAKL 333
RESULT 11
G3P3_ECOLI STANDARD; PRT; 333 AA.
AC P33898; P76094; P78062; P78291; Q03850; Q63208;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase C (EC 1.2.1.12) (GAPDH-C).
GN GAPC OR B1416/B1417.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=River isolate, and Clinical isolate;
RX MEDLINE=98283700; PubMed=9622357;
RA Espinosa-Urgel M., Koller R.;
RT "Escherichia coli genes expressed preferentially in an aquatic
environment."
RL Mol. Microbiol. 28:325-332(1998).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=K12.
RA Hidalgo E., Limon A., Aguilar J.;
RT "A second Escherichia coli gene with similarity to gapA."
RL Microbiology 12:99-106(1996).
RN [3]
SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shaoy Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [4]
SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
RA Sampei G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map."
RL DNA Res. 3:363-377(1996).
RN [5]
SEQUENCE OF 306-332 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92011371; PubMed=1917845;
RA Hidalgo E., Chen Y.-M., Lin E.C.C., Aguilar J.;
RT Molecular cloning and DNA sequencing of the Escherichia coli K-12
ald gene encoding aldehyde dehydrogenase."
RL J. Bacteriol. 173:6118-6123(1991).
RN [6]
SEQUENCE OF 1-13 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=88232416; PubMed=2836696;
RA Nakamura H., Murakami H., Yamato I., Anraku Y.;
RT "Nucleotide sequence of the cybB gene encoding cytochrome b561 in
Escherichia coli K12."
RL Mol. Gen. Genet. 212:1-5(1988).
RN [7]
SEQUENCE OF 1-165 FROM N.A.
RA Krawetz S.A.;
RL Submitted (Oct-1990) to the EMBL/GenBank/DBJ databases.
RN [8]
SEQUENCE OF 1-52 FROM N.A.
RX MEDLINE=87053179; PubMed=3780374;
RA Krawetz S.A., Connor W., Cannon P.D., Dixon G.H.;
RT "A vector-primer-cloner-sequencer plasmid for the construction of cDNA
libraries: evidence for a rat glyceraldehyde-3-phosphate
dehydrogenase-like mRNA and a ferritin mRNA within testis."
RL DNA 5:427-435(1986).
CC -! CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
NAD(+) -> 3-phospho-D-glyceroyl phosphate + NADH.
CC -! PATHWAY: Second phase of glycolysis; first step.
CC -! SUBUNIT: Homotetramer (By similarity).
CC -! SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -! SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
dehydrogenase family.
CC -! CAUTION: IN THE K12 STRAIN THIS GENE IS DISRUPTED BY A STOP CODON
AND A FRAMESHIFT. IT SEEMS TO BE INTACT IN A NUMBER OF WILD
STRAINS.
CC -! CAUTION: Ref.7 and Ref.8 sequences were originally thought to
originate from rat.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
EMBL; L09067; AAA23856.1; ALT_FRAME.
EMBL; A5000239; AAC74498.1; ALT_FRAME.
EMBL; A5000239; AAC74499.1; ALT_FRAME.
EMBL; D30780; BAA15033.1; ALT_FRAME.
EMBL; D30781; BAA15038.1; ALT_FRAME.
EMBL; M64541; -; NOT_ANNOTATED_CDS.
EMBL; X07569; -; NOT_ANNOTATED_CDS.
EMBL; X54798; CAA38569.1; -
EMBL; M4166; AAA11178.1; -
HSSP; P17721; IHGG.
ECcGene; EG12103; gapC.
InterPro; IPR000173; GAP_dhhydrogenase.
InterPro; IPR006424; GAPDH-I.
Pfam; PF000044; gpdh; 1.
```

DR Pfam: PF02800; gpdh.C; 1.
 DR PRINTS: PR00078; G3PDHGRGNASE.
 DR TIGRFAMs: TIGR01534; GAPDH-I; 1.
 DR PROSITE: PS00071; GAPDH; 1.
 KW Glycolysis; Oxidoreductase; NAD; Multigene family; Complete proteome.
 FT BINDING 150 150
 FT ACT_SITE 177 177
 FT CONFLICT 39 39
 FT SEQUENCE 333 AA; 35649 MW; E882223297376BAA0 CRC64;
 Query Match 57.8%; Score 992; DB 1; Length 333;
 Best Local Similarity 59.3%; Pred. No. 1.2e-61;
 Matches 198; Conservative 54; Mismatches 76; Indels 6; Gaps 5;
 QY 4 KVGINGGRIGRLAFRIQNV-GVEVTRINDLTDNMLAHLKYDTTQGRDGTVEVKE 62
 DB 3 KVGINGGRIGRLVGLLELVKSNIDVAINDTSPKILAYLKHDSNYGPPWSDVFT 62
 QY 63 GGFVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVITA 122
 DB 63 DSLIVDGKSTAVYAEKAKNIPWAKAGAEIIVECTGYTSAEKSQAHLDA-GAKKVLISA 121
 QY 123 PGNDVKTVPNTNHDILDTGTTVISGASCTTCLAPMAKALHDAFGIQGLMTTHAYT 182
 DB 122 PAG-EMKTIYVNVNDDTLGNDITVSVASCTTCLAPMAKALHDSFGIEVGTMTTHAYT 180
 QY 183 GDQMLDGGHGGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPTGS 242
 DB 181 GTQSLDGP-RGKDLRASRAAENIIPHTTGAAGAIGLVIPELNGKLDGAAQRPVPTGS 239
 QY 243 VTELVTLDKNVSVDEINAAKAA--SNDGFTYEDPIVSSDIVSVSGSLFATQTKVM 300
 DB 240 VTELVSILGKVTAAEVNNAKQATTNESFGYTDIEIVSSDIIGSHFGSVFATQTEIT 299
 QY 301 EVDGSLVKVSWYDNEMSYTAQLVLTLEYFAKI 334
 DB 300 AVGLQLVKTVAWYDNEYGFVTQIRLTLEFAKL 333
 RESULT 12
 G3P_MYCPN STANDARD; PRT; 337 AA.
 ID G3P_MYCPN
 AC P75358;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
 GN GAPA OR GAP OR MPN430 OR MP411.
 OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 EX MEDLINE=97105885; PubMed=8948633;
 RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
 RA Herrmann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae";
 RL Nucleic Acids Res. 24:4420-4449 (1996).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -1- PATHWAY: Second phase of glycolysis, first step.
 CC -1- SUBUNIT: Homotrimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 CC dehydrogenase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000040; AAB96059.1; -.
 DR PIR: S73737; S73737.
 DR HSSP: P17721; 1HDG.
 DR InterPro: IPR00173; GAP dhhydrogenase.
 DR InterPro: IPR06424; GAPDH-I.
 DR Pfam: PF00044; gpdh.1.
 DR Pfam: PF02800; gpdh.C; 1.
 DR PRINTS: PR00078; G3PDHGRGNASE.
 DR TIGRFAMs: TIGR01534; GAPDH-I; 1.
 DR PROSITE: PS00071; GAPDH; 1.
 KW Glycolysis; Oxidoreductase; NAD; Complete proteome.
 FT BINDING 157 157
 FT ACT_SITE 184 184
 FT SEQUENCE 337 AA; 36805 MW; 550747A529ABCA93 CRC64;
 Query Match 53.7%; Score 921; DB 1; Length 337;
 Best Local Similarity 54.1%; Pred. No. 1e-56;
 Matches 180; Conservative 54; Mismatches 95; Indels 4; Gaps 3;
 QY 3 KVGINGGRIGRLAFRIQNVGVEVTRINDLTDNMLAHLKYDTTQGRDGTVEVKE 62
 DB 8 IRVAINGFGRIGRLVFRALLSQKNIIEIVAVNDLTHPTDLAHLKYDSAHGFKKKVAKD 67
 QY 63 GGFVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVITA 122
 DB 58 NTLMDIKKKLVFSEKDPANIPFAEHNDITVVESTGRFVSEGLSLHQA-GAKEVIISA 126
 QY 123 PG-GNDVKTVPNTNHDILDTGTTVISGASCTTCLAPMAKALHDAFGIQGLMTTHAY 181
 DB 127 PAKQKTITKTVVNVNKHIIAEDKIIISAACTTNCCLAPMVHVLKNGFGLHGTVMVTTHAY 186
 QY 182 TGDQMLDGGHGGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPTG 241
 DB 187 TADQLQDAPH--SDLRARAAACNIVTTTGAAGAIGLVIPEATGKNGVALRVPVLTG 244
 QY 242 SVTELVTLDKNVSVDEINAAKAA--SNDGFTYEDPIVSSDIVSVSGSLFATQTKVM 301
 DB 245 SIVELCVALEKDATVEQINQAKKAASAFRYCEDEIVSSDIVSGSEHGSIFDSKLNTIIE 304
 QY 302 VDGSLVKVSWYDNEMSYTAQLVLTLEYFAKI 334
 DB 305 VDGKLYKYAWYDNESYVNLVVRVNYCAKL 337
 RESULT 13
 G3P_CORGL STANDARD; PRT; 334 AA.
 ID G3P_CORGL
 AC Q01651;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
 GN GAP OR GGL1588.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13059 / AS019;
 RX MEDLINE=93015645; PubMed=1400158;
 RA Eikmanns B.J.;
 RT "Identification, sequence analysis, and expression of a
 RT Corynebacterium glutamicum gene cluster encoding the three glycolytic
 RT enzymes glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate
 RT kinase, and triosephosphate isomerase";


```
RL J. Bacteriol. 174:6076-6086(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RI Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

EMBL; X59403; CAA42045.1; -;
EMBL; AP005279; BAB98981.1; -;
PIR; A43260; A43260.
HSSP; P00362; 1GD1.
InterPro; IPR000173; GAP_dhhydrogenase.
InterPro; IPR006424; GAPDH-I.
Pfam; PF00844; gpdh; 1.
Pfam; PF02800; gpdh_C; 1.
PRINTS; PR00078; G3PDHGRGNASE.
TIGRFAMs; TIGR01534; GAPDH-I; 1.
PROSITE; PS00071; GAPDH; 1.
Glycolysis; Oxidoreductase; NAD; Complete proteome.
BINDING 153 153 GLYCERALDEHYDE 3-PHOSPHATE.
ACT SITE 180 180 ACTIVATES THIOL GROUP DURING CATALYSIS.
CONFLICT 25 26 SD -> NG (IN REF. 1).
CONFLICT 333 334 KL -> QALN (IN REF. 1).
SEQUENCE 334 AA; 36045 MW; 33792AF65FA30FA7 CRC64;

Query Match 53.0%; Score 908.5; DB 1; Length 334;
Best Local Similarity 56.2%; Pred. No. 7.4e-56;
Matches 190; Conservative 41; Mismatches 94; Indels 13; Gaps 6;

QY 1 MVVKGINGFGRIQLAFPRRI-QNVGVEVTRINDLTPNNLAHLKYDTTQGRFDGTV 59
D 1 MTRVINGFGRIQNRFFRAVLERSDDLEVVAVNDLTNDKTLTLTKFDSIMGRLGQVE 60
QY 60 VKEGFEVNGFIKVSARDPENIDWATGDVEIVLEATGFEFAKXAEKHLHANGAKVV 119
D 61 YDDDSITVGGKRIAYARDPKNLDAHNVDIIVESTGFTDANAARAHIEA-GAKVI 119
QY 120 ITAPGNDVKTWFNTNHDILD-GTETVIGASCTTNCPLAPMAKALHDAFGIQKGLMTTI 178
D 120 ISAPASNEDATFVGVGNHESYDPENHNVISGASCTTNCPLAPMAKALHDAFGIQKGLMTTV 179
QY 179 HAYTDQMLDQPHRGGLRPARACANIVSNSTGAKAIGLIVPELNGKLDGAAQRPV 238
D 180 HAYTDQMLDQPHRGGLRPARACANIVSNSTGAKAIGLIVPELNGKLDGAAQRPV 237
QY 239 PTGVSVELVTLIDKNVSYDEINAAKMAASNDISFG----YTBEDPIVSSDIVGSYGLFDA 294
D 238 ITGSATDLTFTNKSEVTVESINAAIKAAVGEFGETLAYSEPLVSTDIVHDSHGSIFDA 297
QY 295 TOTKKNVDSGLVKVSWYDNEMSYTAQIVRTLEYFA 332
D 298 GLTKV-----SGNTVKVSWYDNEMGYTCQLLELTELVA 331

RESULT 14
GAP_MYCGE
ID G3P_MYCGE STANDARD; PRT; 337 AA.

AC P47543;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAPA OR GAP OR MG301.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RA MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischman J.L., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
RN [2]
RP SEQUENCE OF 1-81 AND 279-337 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
RT sequencing.";
RL J. Bacteriol. 175:7918-7930(1993).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotetramer (by similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; U39710; AAC71523.1; -;
EMBL; U02213; AAD12507.1; -;
EMBL; U02178; AAD12463.1; -;
PIR; C64233; C64233.
TIGR; MG301; 1HDG.
InterPro; IPR000173; GAP_dhhydrogenase.
InterPro; IPR006424; GAPDH-I.
Pfam; PF00844; gpdh; 1.
Pfam; PF02800; gpdh_C; 1.
PRINTS; PR00078; G3PDHGRGNASE.
TIGRFAMs; TIGR01534; GAPDH-I; 1.
PROSITE; PS00071; GAPDH; 1.
Glycolysis; Oxidoreductase; NAD; Complete proteome.
BINDING 157 157 GLYCERALDEHYDE 3-PHOSPHATE.
ACT SITE 184 184 ACTIVATES THIOL GROUP DURING CATALYSIS.
FT (BY SIMILARITY).
FT (BY SIMILARITY).
SEQUENCE 337 AA; 37097 MW; FALEA1966687006B CRC64;

Query Match 52.9%; Score 908; DB 1; Length 337;
Best Local Similarity 53.2%; Pred. No. 8.2e-56;
Matches 177; Conservative 54; Mismatches 98; Indels 4; Gaps 3;

QY 3 VKVINGFGRIQLAFPRRIQNVGVEVTRINDLTPNNLAHLKYDTTQGRFDGTV 62
D 8 IKVINGFGRIQLAFPRRIQNVGVEVTRINDLTPNNLAHLKYDTTQGRFDGTV 67

QY	63	GGFEVNGNFIKVSARDPENIDWATGDVGEIVLEATGFFAKKAEAKHLHANGAKKVITA	122
Db	68	NILQIDRKVKVYFSSQPNQLPWFDEHDIDVVIESTGRFVSEEGASLHLKA-GAKRVIIISA	126
QY	123	PG-GNDVKTWVFNTNHDILDGTETVVISGASCTTNCCLAPMAKALHDAFGIOKGLMTTIHAY	181
Db	127	PAKETIRITVVVYVNVNKHKTSSDDKIISAASCTTNCCLAPLVHLEKNFGIVVGTMLTVHAY	186
QY	182	TGDOMILDCGPHRGGLRBARAGAAINVNSTGAAKATGLATPELNGKLDGAAORVBPVTG	241
Db	187	TADQRLODAPH--NDLRRARAANVINVPITTTGAAGA-GLVVVPEANGKLNGLSURVPLVTG	244
QY	242	SVTELVVTLIDKNVSVDEINAAKASNDOSFGYTEDPIVSSDIVGSVGSLSFDATQTKWME	301
Db	245	SIVELSVLVLEKSPSEVQVQNQAKRFASAPKYCEDPIVSSDVSVSSEYGSIFDSKLTNIVE	304
QY	302	VDGSQLVKVSVWYDNEMSVTAQLVRLTLEYFAKI	334
Db	305	VDGMKLYKYAWYDNESYVHQLVRVWSYCAKL	337
RESULT 15			
G3P1_BACSU			
ID	G3P1_BACSU	STANDARD;	PRT; 334 AA.
AC	P09124;		
DT	01-MAR-1989 (Rel. 10, Created)		
DT	01-MAR-1989 (Rel. 10, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Glyceraldehyde 3-phosphate dehydrogenase 1 (EC 1.2.1.12) (GAPDH) (NAD-		
DE	dependent glyceraldehyde-3-phosphate dehydrogenase).		
GN	GAPA OR GAP OR BSU33940.		
OS	Bacillus subtilis.		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
OX	NCBI_TaxID=1423;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=168 / BD170;		
PX	MEDLINE=89160255; PubMed=2493629;		
RA	Vlaene A., Dhese P.;		
RT	"Sequence of the glyceraldehyde-3-phosphate dehydrogenase gene from		
RT	Bacillus subtilis."		
RL	Nucleic Acids Res. 17:1251-1251 (1989).		

```

RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RL subtilis.";
RN Nature 390:249-256(1997).
RC [3]
RP SEQUENCE OF 1-30.
RC STRAIN=166 / JH642;
RX MEDLINE=96345629; PubMed=8755892;
RA Graumann P., Schroeder K., Schmid R., Marahiel M.A.;
RT "Cold shock stress-induced proteins in Bacillus subtilis.";
RN J. Bacteriol. 178:4611-4619(1996).
RC [4]
RP CHARACTERIZATION.
RX MEDLINE=20261518; PubMed=10799476;
RA Fillinger S., Boschi-Nuller S., Azza S., Dervyn E., Branlant G.,
RA Aymerich S.;
RT "Two glyceraldehyde-3-phosphate dehydrogenases with opposite
RL physiological roles in a nonphotosynthetic bacterium.";
RN J. Biol. Chem. 275:14031-14037(2000).
CC -1- FUNCTION: More active in catabolism.
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X13011; CA931434.1; -.
DR EMBL; Z99121; CAB15399.1; -.
DR PIR; S02754; DEBSG.
DR HSP; P00362; IGD1.
DR Subtilist; BG10827; gapA.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR InterPro; IPR006424; GAPDH-1.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; 1.
DR PRINTS; PR00078; G3PDHGRNASE.
DR TIGRfams; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD; Multigene family; Complete proteome.
FT INIT MET 0
FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE.
FT ACT SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ SEQUENCE 334 AA; 35701 MW; 1283D3B6CF5095EC CRC64;

Query Match 52.0%; Score 891.5; DB 1; Length 334;
Best Local Similarity 54.8%; Pred.No.1.le-54;
Matches 183; Conservative 41; Mismatches 101; Indels 11; Gaps 5;

Qy 3 VKYINGFGFRIGRLAFRRIQNVGVETVTRINDLPDNLHLKLYDTQGRFDGTVEVKE 62
Db 2 VKYINGFGFRIGENVFRALNNPEVVAVDLTDANMLAHLQYDSVHGKDAEVSVDG 61
Qy 63 GGFVNGNFKVISAERDPENIDWATGVETVIEATGFFAKKZAAEKLHANGAKKVVITA 122
Db 62 NNIVVNGKTIIEVAERDPAKLSGKGQVETVVESTGFTTKRADAAKHLEA-GAKKVIISA 120
Qy 123 PGGNDVKTVVFNTHDILD-GTETVTSASCTTNCLAPMAKALHDAFGIQKGLMTTHAY 181
Db 121 PANEEDITIVMGVNEKYDAANDHDVINSASCTTNCLAPKVLNDRKFGIKRGMTTVHSY 180
Qy 182 TGDQMTLDGPHRGDGLRRARAGAAATVPNASTGAAGAIGLIVIPELNGKLDGAARVPVPTG 241
Db 181 TNDQOILDLPKH--DYRRARAANAETIPSTGAAKAVSLVLPKLGKLGNGAMRVPTNV 238

```

	Qy	242	SYTELVTLDKNVSVDINRAKKAASND-----FGYTEDPIVSSDIVGSYSGSLPDAQTQ	297
	Db	239	SUVDVAELNEGVTAEEVNAALKAARGDGLKGILGHSSEETNGNKNGSTIDALST	298
	Qy	298	KMVEVDGSQLVKVSVSWTDNMESYTAQLVRTLEYFAK	333
	Db	299	MYME---GSMVKVISWTDNBEGSYGNRVVDAAVIAK	331

Search completed: May 11, 2004, 16:48:17
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 16:45:16 ; Search time 46 Seconds
(without alignments)
2304.655 Million cell updates/sec

Title: US-10-650-369-12
Perfect score: 1715
Sequence: 1 MVVKVINGFGRIGRLAFRR.....EMSYTAQLVRLTEYFAKIAK 336

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_protein:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1715	100.0	336	2 Q83ZF7	Q83zf7 streptococc
2	1618	94.3	336	2 Q8KVU6	Q8kvu6 streptococc
3	1609	93.8	336	2 Q9ALW2	Q9alw2 streptococc
4	1609	93.8	336	16 Q8E3E8	Q8e3e8 streptococc
5	1609	93.8	336	16 Q8XS8	Q8xs8 streptococc
6	1596	93.1	336	2 Q8GCR7	Q8gcr7 streptococc
7	1593.5	92.9	335	16 Q97NL1	Q97nl1 streptococc
8	1593.5	92.9	335	16 Q8CWN6	Q8cwn6 streptococc
9	1584	92.4	336	2 Q8KHG1	Q8khg1 streptococc
10	1579.5	92.1	337	16 Q8DVV3	Q8dvv3 streptococc
11	1578	92.0	336	2 Q8KVU5	Q8kvu5 streptococc
12	1572	91.7	336	2 Q8VVB9	Q8vvb9 streptococc
13	1518	88.5	320	2 Q9L5X6	Q9l5x6 streptococc
14	1468.5	85.6	309	2 Q9AJT7	Q9ajt7 streptococc
15	1467.5	85.6	309	2 Q9AJT9	Q9ajt9 streptococc
16	1465.5	85.5	309	2 Q9AJT4	Q9ajt4 streptococc

17 1463.5 85.3 309 2 Q9AJT5
18 1462.5 85.3 308 2 Q9AJT8
19 1445.5 84.3 305 2 Q9AJT6
20 1410 82.2 336 16 Q9CDH4
21 1337.5 78.0 333 16 Q83318
22 1262 73.6 336 16 Q8V411
23 1261 73.5 336 16 Q928H9
24 1227.5 71.6 332 16 Q8XKT9
25 1218.5 71.0 335 16 Q8RFN9
26 1197.5 69.8 334 16 Q9JW78
27 1190 69.4 336 16 Q8CFY5
28 1189.5 69.4 330 2 Q83UX3
29 1186.5 69.2 330 2 Q84H26
30 1185.5 69.1 330 2 Q84H24
31 1184.5 69.1 334 16 Q9UX95
32 1183 69.0 336 16 Q925C5
33 1182.5 69.0 330 2 Q84H25
34 1178.5 68.7 335 2 Q93M61
35 1138 66.4 311 2 Q8L2R0
36 1132.5 66.0 310 2 Q8L2P7
37 1124.5 65.6 310 2 Q8L2E6
38 1124 65.5 311 2 Q8L2P8
39 1123.5 65.5 310 2 Q8L2Q6
40 1121.5 65.4 311 2 Q8VM79
41 1120.5 65.3 310 2 Q8L2Q3
42 1120.5 65.3 310 2 Q8L2P5
43 1118.5 65.2 310 2 Q8L2Q2
44 1116.5 65.1 310 2 Q8L2R7
45 1115.5 65.0 310 2 Q8L2R2

ALIGNMENTS

RESULT 1
Q83ZF7 ID Q83ZF7 PRELIMINARY; PRT; 336 AA.
AC Q83ZF7;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Glyceraldehyde 3-P dehydrogenase.
GN GAPC.
OS Streptococcus dysgalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1334;
RN [1]
RP SEQUENCE FROM N.A.
RA Perez-Casal J.F., Bolton A., Song X.-M., Willson P., Potter A.A.;
RT "Use of the surface proteins GapC and Mig of Streptococcus
RT dysgalactiae as protective antigens against mastitis in non-lactating
RT cows";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF375662; AAP31408.1;
DR GO; GO:0004385; F:glyceraldehyde-3-phosphate dehydrogenase (p. . . ; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR00173; GAP_dhdrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh.C; 1.
DR PRINTS; PR00078; G3PDHGRGNASE.
DR TIGRPFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
SQ SEQUENCE 336 AA; 35928 MW; 11828218CF037076 CRC64;

Query Match 100.0%; Score 1715; DB 2; Length 336;
Best Local Similarity 100.0%; Pred. No. 5.7e-107;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLINDLPNMLAHLLKYDTTQGRPDGVEV 60
DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLINDLPNMLAHLLKYDTTQGRPDGVEV 60

Q9ajt5 streptococc
Q9ajt8 streptococc
Q9ajt6 streptococc
Q9cdh4 lactococcu
Q83318 enterococc
Q8v411 listeria m
Q928h9 listeria in
Q8xkt9 clostridium
Q8rfn9 fusobacteri
Q9jw78 neisseria m
Q8cfy5 staphylococ
Q83ux3 neisseria g
Q84h26 neisseria g
Q84h24 neisseria g
Q9ux95 neisseria m
Q925c5 staphylococ
Q84h25 neisseria g
Q93m61 staphylococ
Q8l2r0 staphylococ
Q8l2p7 staphylococ
Q8l2p6 staphylococ
Q8l2p8 staphylococ
Q8l2q6 staphylococ
Q8vm79 rhodospiril
Q8l2q3 staphylococ
Q8l2p5 staphylococ
Q8l2q2 staphylococ
Q8l2r7 staphylococ
Q8l2r2 staphylococ

QY	61	KEGGEVNGNFIKVSARDPENIDWATDGVIEIATGFFAKKAAEKLHANGAKKVI	120
Db	61	KEGGEVNGNFIKVSARDPENIDWATDGVIEIATGFFAKKAAEKLHANGAKKVI	120
QY	121	TAPGNDVKTVFNTHDILDTETVSGASCTTNCLAPMAKALHDAFGIQKGLMTTHA	180
Db	121	TAPGNDVKTVFNTHDILDTETVSGASCTTNCLAPMAKALHDAFGIQKGLMTTHA	180
QY	181	YTDQMILDPHRRAGDGLRRRAGAAANIPNSTGAAGAIGLVIPELNGKLDGAARVPVPT	240
Db	181	YTDQMILDPHRRAGDGLRRRAGAAANIPNSTGAAGAIGLVIPELNGKLDGAARVPVPT	240
QY	241	GSVTLEWTLTKNVSDEINAAKAAANDSFGYTEDPIVSSDIVGVSGLFDATQTKVM	300
Db	241	GSVTLEWTLTKNVSDEINAAKAAANDSFGYTEDPIVSSDIVGVSGLFDATQTKVM	300
QY	301	EVDGSQLVKVSVDNEMSYTAQLVRLTLEYFAKIAK	336
Db	301	EVDGSQLVKVSVDNEMSYTAQLVRLTLEYFAKIAK	336
RESULT 2			
Q8KVU6			
ID	Q8KVU6	PRELIMINARY; PRT; 336 AA.	
AC	Q8KVU6		
DT	01-OCT-2002 (TremBLrel. 22, Created)		
DT	01-OCT-2002 (TremBLrel. 22, Last sequence update)		
DT	01-OCT-2003 (TremBLrel. 25, Last annotation update)		
DE	Glyceraldehyde-3-phosphate dehydrogenase.		
OS	Streptococcus uberis.		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
CC	Streptococcus.		
OX	NCBI_TaxID=1349;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ATCC 9297;		
RA	Fontaine M.C., Perez-Casal J., Song X.-M., Shelford J., Willson P.J.,		
RA	Potter A.A.;		
RT	"Immunization of dairy cattle with recombinant GapC and chimeric CAMP		
RT	antigens confers protection against heterologous challenge with		
RT	Streptococcus uberis."		
RL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.		
CC	-!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE		
CC	DEHYDROGENASE FAMILY.		
DR	EMBL; AF421900; AAM73771.1; ..		
DR	GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . . ; IEA.		
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.		
DR	GO; GO:0006096; P:glycolysis; IEA.		
DR	InterPro; IPR006424; GAPDH-I.		
DR	InterPro; IPR00173; GAP_dhydrogenase.		
DR	Pfam; PF00044; gpdh; 1.		
DR	Pfam; PF02800; gpdh; 1.		
DR	PRINTS; PR00078; G3PDHGRGNASE.		
DR	TIGRFAMs; TIGR01534; GAPDH-I; 1.		
DR	PROSITE; PS00071; GAPDH; 1.		
KW	Oxidoreductase.		
SQ	SEQUENCE 336 AA; 35906 MW; 755D74A4548B61D4 CRC64;		
Query Match 94.3%; Score 1618; DB 2; Length 336;			
Best Local Similarity 92.9%; Pred. No. 1.8e-100;			
Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;			
QY	1	MVVKVINGFGFRIGLAFRRQNVGEVTRINDLTDPMLAHLKYDTTQGRFDGTVVEV	60
Db	1	MVVKVINGFGFRIGLAFRRQNVGEVTRINDLTDPMLAHLKYDTTQGRFDGTVVEV	60
QY	61	KEGGEVNGNFIKVSARDPENIDWATDGVIEIATGFFAKKAAEKLHANGAKKVI	120
Db	61	KEGGEVNGNFIKVSARDPENIDWATDGVIEIATGFFAKKAAEKLHANGAKKVI	120
QY	121	TAPGNDVKTVFNTHDILDTETVSGASCTTNCLAPMAKALHDAFGIQKGLMTTHA	180
Db	121	TAPGNDVKTVFNTHDILDTETVSGASCTTNCLAPMAKALHDAFGIQKGLMTTHA	180

Db	121	TAPGDDVKTVFNTHDILDTETVSGASCTTNCLAPMAKALQDNFVGKQGLMTTHA	180
QY	181	YTDQMILDPHRRAGDGLRRRAGAAANIPNSTGAAGAIGLVIPELNGKLDGAARVPVPT	240
Db	181	YTDQMILDPHRRAGDGLRRRAGAAANIPNSTGAAGAIGLVIPELNGKLDGAARVPVPT	240
QY	241	GSVTLEWTLTKNVSDEINAAKAAANDSFGYTEDPIVSSDIVGVSGLFDATQTKVM	300
Db	241	GSVTLEWTLTKNVSDEINAAKAAANDSFGYTEDPIVSSDIVGVSGLFDATQTKVM	300
QY	301	EVDGSQLVKVSVDNEMSYTAQLVRLTLEYFAKIAK	336
Db	301	EVDGSQLVKVSVDNEMSYTAQLVRLTLEYFAKIAK	336
RESULT 3			
Q9ALW2			
ID	Q9ALW2	PRELIMINARY; PRT; 336 AA.	
AC	Q9ALW2		
DT	01-JUN-2001 (TremBLrel. 17, Created)		
DT	01-JUN-2001 (TremBLrel. 17, Last sequence update)		
DT	01-OCT-2003 (TremBLrel. 25, Last annotation update)		
DE	Glyceraldehyde 3-phosphate dehydrogenase.		
OS	Streptococcus agalactiae.		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
CC	Streptococcus.		
OX	NCBI_TaxID=1311;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=J48;		
RA	Seifert K.N., Bleiweis A.S., McArthur W.P., Brady L.J.;		
RT	"The Group B Streptococcal Surface Antigen Delta is a Glyceraldehyde		
RT	3-Phosphate Dehydrogenase."		
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.		
CC	-!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE		
CC	DEHYDROGENASE FAMILY.		
DR	EMBL; AF338416; AAK14387.1; ..		
DR	HSSP; P00362; 1GDI		
DR	GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . . ; IEA.		
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.		
DR	GO; GO:0006096; P:glycolysis; IEA.		
DR	InterPro; IPR006424; GAPDH-I.		
DR	InterPro; IPR00173; GAP_dhydrogenase.		
DR	Pfam; PF00044; gpdh; 1.		
DR	Pfam; PF02800; gpdh; 1.		
DR	PRINTS; PR00078; G3PDHGRGNASE.		
DR	TIGRFAMs; TIGR01534; GAPDH-I; 1.		
DR	PROSITE; PS00071; GAPDH; 1.		
KW	Oxidoreductase.		
SQ	SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;		
Query Match 93.8%; Score 1609; DB 2; Length 336;			
Best Local Similarity 92.0%; Pred. No. 7.2e-100;			
Matches 309; Conservative 17; Mismatches 10; Indels 0; Gaps 0;			
QY	1	MVVKVINGFGFRIGLAFRRQNVGEVTRINDLTDPMLAHLKYDTTQGRFDGTVVEV	60
Db	1	MVVKVINGFGFRIGLAFRRQNVGEVTRINDLTDPMLAHLKYDTTQGRFDGTVVEV	60
QY	61	KEGGEVNGNFIKVSARDPENIDWATDGVIEIATGFFAKKAAEKLHANGAKKVI	120
Db	61	KEGGEVNGNFIKVSARDPENIDWATDGVIEIATGFFAKKAAEKLHANGAKKVI	120
QY	121	TAPGNDVKTVFNTHDILDTETVSGASCTTNCLAPMAKALHDAFGIQKGLMTTHA	180
Db	121	TAPGNDVKTVFNTHDILDTETVSGASCTTNCLAPMAKALHDAFGIQKGLMTTHA	180
QY	181	YTDQMILDPHRRAGDGLRRRAGAAANIPNSTGAAGAIGLVIPELNGKLDGAARVPVPT	240
Db	181	YTDQMILDPHRRAGDGLRRRAGAAANIPNSTGAAGAIGLVIPELNGKLDGAARVPVPT	240
QY	241	GSVTLEWTLTKNVSDEINAAKAAANDSFGYTEDPIVSSDIVGVSGLFDATQTKVM	300
Db	241	GSVTLEWTLTKNVSDEINAAKAAANDSFGYTEDPIVSSDIVGVSGLFDATQTKVM	300

Db 241 GSVTELVATLEKDVTVVEVNAAMKAAANDSYGTYEDPIVSSDIVGSGSLFDATQTKVQ 300
Qy 301 EVDGSQLVKVSWYDNEMSYTAQLVLTLEYFAKIAK 336
Db 301 TVDGNQLVKVSWYDNEMSYTSLVLTLEYFAKIAK 336
RESULT 4
Q8E3E8 PRELIMINARY; PRT; 336 AA.
ID Q8E3E8
AC Q8E3E8
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase.
GN GBS1811.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEM316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,
RA Meadek T., Zouine M., Couve E., Lailoui L., Foyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease."
RL Mol. Microbiol. 45:1499-1513 (2002).
DR EMBL; AL766853; CAD47470.1; -.
DR Sgallist; gbs1811; -.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; 1.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PRINTS; PR00078; G3PDHGRNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Complete proteome.
SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

Query Match 93.8%; Score 1609; DB 16; Length 336;
Best Local Similarity 92.0%; Pred. No. 7.2e-100;
Matches 309; Conservative 17; Mismatches 10; Indels 0; Gaps 0;
Qy 1 MVVKVINGFGRIQRLAFRRIONVEGVETVTRINDLTPNNLAHLKLYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIQRLAFRRIONVEGVETVTRINDLTPNNLAHLKLYDTTQGRFDGTVEV 60
Qy 61 KEGGFVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120
Db 61 KEGGFVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120
Qy 121 TAPGNDVKTVPNTNHDILDTETVIGSCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 180
Db 121 TAPGNDVKTVPNTNHDILDTETVIGSCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 180
Qy 181 YTGQNDVKTVPNTNHDILDTETVIGSCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 240
Db 181 YTGQNDVKTVPNTNHDILDTETVIGSCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 240
Qy 241 GSVTELVATLEKDVTVVEVNAAMKAAANDSYGTYEDPIVSSDIVGSGSLFDATQTKVQ 300
Db 241 GSVTELVATLEKDVTVVEVNAAMKAAANDSYGTYEDPIVSSDIVGSGSLFDATQTKVQ 300
Qy 301 EVDGSQLVKVSWYDNEMSYTAQLVLTLEYFAKIAK 336
Db 301 TVDGNQLVKVSWYDNEMSYTSLVLTLEYFAKIAK 336

RESULT 6

RESULT 5
Q8DXS8 PRELIMINARY; PRT; 336 AA.
ID Q8DXS8
AC Q8DXS8
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase.
GN GAP OR SAG1768.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=2222988; PubMed=12200547;
RA Tettelin H., Maignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margalit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radu D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
RA Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).
DR EMBL; AE014272; AAN00631.1; -.
DR TIGR; SAG1768; -.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; 1.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PRINTS; PR00078; G3PDHGRNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Complete proteome.
SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

Query Match 93.8%; Score 1609; DB 16; Length 336;
Best Local Similarity 92.0%; Pred. No. 7.2e-100;
Matches 309; Conservative 17; Mismatches 10; Indels 0; Gaps 0;
Qy 1 MVVKVINGFGRIQRLAFRRIONVEGVETVTRINDLTPNNLAHLKLYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIQRLAFRRIONVEGVETVTRINDLTPNNLAHLKLYDTTQGRFDGTVEV 60
Qy 61 KEGGFVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120
Db 61 KEGGFVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAEKHLHANGAKKVI 120
Qy 121 TAPGNDVKTVPNTNHDILDTETVIGSCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 180
Db 121 TAPGNDVKTVPNTNHDILDTETVIGSCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 180
Qy 181 YTGQNDVKTVPNTNHDILDTETVIGSCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 240
Db 181 YTGQNDVKTVPNTNHDILDTETVIGSCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 240
Qy 241 GSVTELVATLEKDVTVVEVNAAMKAAANDSYGTYEDPIVSSDIVGSGSLFDATQTKVQ 300
Db 241 GSVTELVATLEKDVTVVEVNAAMKAAANDSYGTYEDPIVSSDIVGSGSLFDATQTKVQ 300
Qy 301 EVDGSQLVKVSWYDNEMSYTAQLVLTLEYFAKIAK 336
Db 301 TVDGNQLVKVSWYDNEMSYTSLVLTLEYFAKIAK 336

مجلس

DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (Phosphorylating)
 DE (EC 1.2.1.12).
 GN GAPC OR SPRI825.
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=171101;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burtgett S.,
 RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 RA Gilmour R., Glas J.S., Khoja H., Kraft A.R., Lagace R.E.,
 RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
 RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
 RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,
 RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostock P.R. Jr., Skatrud P.L.,
 RA Glass J.I.;
 RL "Genome of the bacterium Streptococcus pneumoniae strain R6";
 DR J. Bacteriol. 183:5709-5717 (2001).
 DR EMBL; AE008547; AAL00428.1; -.
 DR PIR; G98099; G98099.
 DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR006424; GAPDH-I.
 DR InterPro; IPR00173; GAP_dhdrogenase.
 DR Pfam; PF00044; gpdh; 1.
 DR Pfam; PF02800; gpdh_C; 1.
 DR PRINTS; PR00078; G3PDHGRGNASE.
 DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 DR Oxidoreductase; Complete proteome.
 KW Oxidoreductase.
 SQ SEQUENCE 359 AA; 38763 MW; 61EFOE375B330B60 CRC64;

 Query Match 92.9%; Score 1593.5; DB 15; Length 359;
 Best Local Similarity 92.6%; Pred. No. 8.7e-99;
 Matches 311; Conservative 12; Mismatches 12; Indels 1; Gaps 1;

 QY 1 MVVKVINGFGRIQLAFRIQNVGEVETRLNDLTPNMLHLKYDTTQGRFDGTVEV 60
 DB 25 MVVKVINGFGRIQLAFRIQNVGEVETRLNDLTPNMLHLKYDTTQGRFDGTVEV 84
 QY 61 KEGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAEKHLHANGAKKVI 120
 DB 85 KEGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAEKHLHANGAKKVI 143
 QY 121 TAPGNDVKTVVNTNHDLDGTETVSGASCTTNCLAPMAKALHDAFGIOKGLMTTIIHA 180
 DB 144 TAPGNDVKTVVNTNHDLDGTETVSGASCTTNCLAPMAKALQDNFVGKGLMTTIIHA 203
 QY 181 YTGDMILDPGHRGDLRRARAGAAANIPVNSTGAAGAIGLVIPELNGKLDGAORVPVPT 240
 DB 204 YTGDMILDPGHRGDLRRARAGAAANIPVNSTGAAGAIGLVIPELNGKLDGAORVPVPT 263
 QY 241 GSVELVAVLEKNTVDVNAAMKAANDSFYTEDPIVSSDIVGSGSLFDATQTKVM 300
 DB 264 GSVELVAVLEKNTVDVNAAMKAANDSFYTEDPIVSSDIVGSGSLFDATQTKVL 323
 QY 301 EVDGSQLVKVVSWYDNEMSYTAQLVRTLEYFAKIAK 336
 DB 324 DVDGQVLKVVSWYDNEMSYTAQLVRTLEYFAKIAK 359

 RESULT 9
 Q8KHG1 ID Q8KHG1 PRELIMINARY; PRT; 336 AA.
 AC Q8KHG1
 DT 01-OCT-2002 (Tremblrel. 22, Created)

DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 OS Streptococcus agalactiae, and
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1311, 1346;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA SPECIES=S agalactiae, and S. pneumoniae; STRAIN=ATCC 27541, and 9117;
 RA Fontaine M.C., Perez-Casal J., Song X.-M., Sheldford J., Willson P.J.,
 RA Potter A.A.;
 RT "Immunization of dairy cattle with recombinant GapC and chimeric CAMP
 RT antigens confers protection against heterologous challenge with
 RT Streptococcus uberis";
 RL Streptococcus uberis.
 CC Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 DR EMBL; AF421899; AAM73770.1; -.
 DR EMBL; AF421902; AAM73773.1; -.
 DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR006424; GAPDH-I.
 DR InterPro; IPR00173; GAP_dhdrogenase.
 DR Pfam; PF00044; gpdh; 1.
 DR Pfam; PF02800; gpdh_C; 1.
 DR PRINTS; PR00078; G3PDHGRGNASE.
 DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 DR Oxidoreductase.
 KW Oxidoreductase.
 SQ SEQUENCE 336 AA; 35723 MW; ABAB9E14F3ED1111 CRC64;

 Query Match 92.4%; Score 1584; DB 2; Length 336;
 Best Local Similarity 90.5%; Pred. No. 3.4e-98;
 Matches 304; Conservative 19; Mismatches 13; Indels 0; Gaps 0;

 QY 1 MVVKVINGFGRIQLAFRIQNVGEVETRLNDLTPNMLHLKYDTTQGRFDGTVEV 60
 DB 1 MVVKVINGFGRIQLAFRIQNVGEVETRLNDLTPNMLHLKYDTTQGRFDGTVEV 60
 QY 61 KEGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAEKHLHANGAKKVI 120
 DB 61 KEGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAEKHLHANGAKKVI 120
 QY 121 TAPGNDVKTVVNTNHDLDGTETVSGASCTTNCLAPMAKALHDAFGIOKGLMTTIIHA 180
 DB 121 TAPGNDVKTVVNTNHDLDGTETVSGASCTTNCLAPMAKALQDNFVGKGLMTTIIHG 180
 QY 181 YTGDMILDPGHRGDLRRARAGAAANIPVNSTGAAGAIGLVIPELNGKLDGAORVPVPT 240
 DB 181 YTGDMILDPGHRGDLRRARAGAAANIPVNSTGAAGAIGLVIPELNGKLDGAORVPVPT 240
 QY 241 GSVELVAVLEKNTVDVNAAMKAANDSFYTEDPIVSSDIVGSGSLFDATQTKVM 300
 DB 241 GSVELVAVLEKNTVDVNAAMKAANDSFYTEDPIVSSDIVGSGSLFDATQTKVQ 300
 QY 301 EVDGSQLVKVVSWYDNEMSYTAQLVRTLEYFAKIAK 336
 DB 301 TVDGNQLKVVSWYDNEMSYTAQLVRTLEYFAKIAK 336

 RESULT 10
 Q8DVV3 ID Q8DVV3 PRELIMINARY; PRT; 337 AA.
 AC Q8DVV3
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Extracellular glyceraldehyde-3-phosphate dehydrogenase (EC
 DE 1.2.1.12).
 GN GAPC OR SMU.360.


```

OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Jin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
RT pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
RL EMBL; AE014893; AAN58118.1; -.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhhydrogenase.
DR Pfam; PF00044; gpdh; 1.
DR PRINTS; PR00078; G3PDHGRGNASE.
DR TIGRfam; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
DR OXidoreductase; Complete proteome.
KW SEQUENCE 337 AA; 36068 MW; 42BF20365963C22 CRC64;

Query Match 92.1%; Score 1579.5; DB 16; Length 337;
Best Local Similarity 90.2%; Pred. No. 6.9e-98;
Matches 304; Conservative 18; Mismatches 14; Indels 1; Gaps 1;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLDTPNMLAHLKYDTTQGRFDGTVEV 60
DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLDTPNMLAHLKYDTTQGRFDGTVEV 60

QY 61 KEGGFVNGNFIKVSARDPENIDWATGVIVLEATGFFAKKEAAEKHLHANGAKKVV 119
DB 61 KEGGFVNGNFIKVSARDPENIDWATGVIVLEATGFFAKKEAAEKHLHANGAKKVV 120

QY 120 ITAPGNDVKTIVNTNHHILDGTETVIGASCTTNCPLAPMAKALHDAFGIQKGLMTTII 179
DB 120 ITAPGNDVKTIVNTNHHILDGTETVIGASCTTNCPLAPMAKALHDAFGIQKGLMTTII 180

QY 180 AYTGDQMLDGPGRGDLRRARAGANNIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVP 239
DB 180 AYTGDQMLDGPGRGDLRRARAGANNIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVP 240

QY 240 TGSVTELVVTLDKVNSVDENAAKKAASNDISFGYTEDPIVSSDIVGVSGLFDATQTKV 299
DB 240 TGSVTELVVTLDKVNSVDENAAKKAASNDISFGYTEDPIVSSDIVGVSGLFDATQTKV 300

QY 300 MEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
DB 301 LDVDGKQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 337

RESULT 11
Q8KVU5 PRELIMINARY; PRT; 336 AA.
ID Q8KVU5;
AC Q8KVU5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
OS Streptococcus parauberis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1348;
RN [1]
RP SEQUENCE FROM N.A.
RA Fontaine M.C., Perez-Casas J., Song X.-M., Shelford J., Willson P.J.,
RA Potter A.A.;

"Immunization of dairy cattle with recombinant GapC and chimeric CAMP
antigens confers protection against heterologous challenge with
Streptococcus uberis.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF421901; AAM73772.1; -.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhhydrogenase.
DR Pfam; PF00044; gpdh; 1.
DR PRINTS; PR00078; G3PDHGRGNASE.
DR TIGRfam; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
DR OXidoreductase.
KW SEQUENCE 336 AA; 36037 MW; 311647C25489AC9E CRC64;

Query Match 92.0%; Score 1578; DB 2; Length 336;
Best Local Similarity 90.8%; Pred. No. 8.6e-98;
Matches 305; Conservative 16; Mismatches 15; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLDTPNMLAHLKYDTTQGRFDGTVEV 60
DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLDTPNMLAHLKYDTTQGRFDGTVEV 60

QY 61 KEGGFVNGNFIKVSARDPENIDWATGVIVLEATGFFAKKEAAEKHLHANGAKKVV 120
DB 61 KEGGFVNGNFIKVSARDPENIDWATGVIVLEATGFFAKKEAAEKHLHANGAKKVV 120

QY 121 TAPGNDVKTIVNTNHHILDGTETVIGASCTTNCPLAPMAKALHDAFGIQKGLMTTII 180
DB 121 TAPGNDVKTIVNTNHHILDGTETVIGASCTTNCPLAPMAKALHDAFGIQKGLMTTII 180

QY 181 YTGDMQLDGPGRGDLRRARAGANNIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVP 240
DB 181 YTGDMQLDGPGRGDLRRARAGANNIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVP 240

QY 241 GSVTELVVTLDKVNSVDENAAKKAASNDISFGYTEDPIVSSDIVGVSGLFDATQTKV 300
DB 241 GSVTELVVTLDKVNSVDENAAKKAASNDISFGYTEDPIVSSDIVGVSGLFDATQTKV 300

QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
DB 301 TVDGNQLVKVSWYDNEMSYTAQLDRLTLEYFAKIAK 336

RESULT 12
Q8VVB9 PRELIMINARY; PRT; 336 AA.
ID Q8VVB9;
AC Q8VVB9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
GN GAPDH.
OS Streptococcus thermophilus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1308;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LMG18311;
RA van den Bogaard P.T.C., Kleerebezem M., Hols P., Crispie F.,
RA Kuipers O.P., de Vos W.M.;
RT "Modulation of Glycolysis by Lactose Availability in Streptococcus
RT thermophilus."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF442551; AAL35377.1; -.

```

DR HSP; P00354; 30PD.
 DR GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
 DR GO:0016491; F:oxidoreductase activity; IEA.
 DR GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR006424; GAPDH-I.
 DR InterPro: IPR000173; GAP_dhrogenase.
 DR Pfam: PF00044; gpdh; 1.
 DR Pfam: PF02800; gpdh; C; 1.
 DR PRINTS; PR00078; G3PDHGRNASE.
 DR TIGRfams; TIGR01534; GAPDH-I; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 DR KW Oxidoreductase.
 DR FT NON TER 320 320
 DR SQ SEQUENCE 336 AA; 36026 MW; 52C1F25F3A7E0230 CRC64;
 Query Match 91.7%; Score 1572; DB 2; Length 336;
 Best Local Similarity 90.2%; Pred. No. 2.2e-97;
 Matches 303; Conservative 18; Mismatches 15; Indels 0; Gaps 0;
 QY 1 MVKVGNGFGRIGRLAFRIQNVGEVETRLNDLTPNMLHLKYDTTQGRFDGTVEV 60
 Db 1 MVKVGNGFGRIGRLAFRIQNVGEVETRLNDLTPNMLHLKYDTTQGRFDGTVEV 60
 QY 61 KEGFVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKAEKHLHANGAKKVI 120
 Db 61 KEGFVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKAEKHLHANGAKKVI 120
 QY 121 TAPGGNDVKTIVFNTNHDILDTGTFVTSASCTTNCPLAPMAKALHDAFCI 180
 Db 121 TAPGGNDVKTIVFNTNHDILDTGTFVTSASCTTNCPLAPMAKALHDAFCI 180
 QY 181 YTGDMILDPGPHRGDLRRARAGAAINVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
 Db 181 YTGDMILDPGPHRGDLRRARAGAAINVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPT 240
 QY 241 GSVTELVTLKNSVDEINAAKASNDSPFGTDEPIVSSDIVGVSGLSDATQTKVM 300
 Db 241 GSVTELVTLKNSVDEINAAKASNDSPFGTDEPIVSSDIVGVSGLSDATQTKVM 300
 QY 301 EVDGSLVKVSVYDNEMSYTAQLVRLTFYFAKIAK 336
 Db 301 EVDGSLVKVSVYDNEMSYTAQLVRLTFYFAKIAK 336
 RESULT 13
 ID Q9L5X6 PRELIMINARY; PRT; 320 AA.
 AC Q9L5X6;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Extracellular glyceraldehyde-3-phosphate dehydrogenase
 DE (Fragment).
 OS Streptococcus gordonii.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 CX NCBI_TaxID=1302;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nelson D., Goldstein J.M., Boatright K., Hartley D.W.S., Cook S.L.,
 RA Hickman P.J., Potempa J., Travis J., Mayo J.A.;
 RT "Purification and characterization of an extracellular
 RT glyceraldehyde-3-phosphate dehydrogenase from Streptococcus sanguis
 RT and cloning of the gene encoding this enzyme."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 DR EMBL; AF247678; AAF64063.1; -.
 DR HSP; P00362; 1GDI.
 DR GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
 DR GO:0016491; F:oxidoreductase activity; IEA.
 DR GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR006424; GAPDH-I.
 DR InterPro: IPR000173; GAP_dhrogenase.

DR Pfam: PF00044; gpdh; 1.
 DR Pfam: PF02800; gpdh; C; 1.
 DR PRINTS; PR00078; G3PDHGRNASE.
 DR TIGRfams; TIGR01534; GAPDH-I; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 DR KW Oxidoreductase.
 DR FT NON TER 320 320
 DR SQ SEQUENCE 320 AA; 33968 MW; 6FFC18BB01E91F0 CRC64;
 Query Match 88.5%; Score 1518; DB 2; Length 320;
 Best Local Similarity 91.2%; Pred. No. 8.4e-94;
 Matches 292; Conservative 14; Mismatches 14; Indels 0; Gaps 0;
 QY 4 KVGINGFGRIGRLAFRIQNVGEVETRLNDLTPNMLHLKYDTTQGRFDGTVEVKEG 63
 Db 1 KVGINGFGRIGRLAFRIQNVGEVETRLNDLTPNMLHLKYDTTQGRFDGTVEVKEG 60
 QY 64 GFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKAEKHLHANGAKKVI 123
 Db 61 GFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKAEKHLHANGAKKVI 120
 QY 124 GGNVDKTVFNTNHDILDTGTFVTSASCTTNCPLAPMAKALHDAFCI 183
 Db 121 GGNVDKTVFNTNHDILDTGTFVTSASCTTNCPLAPMAKALHDAFCI 180
 QY 184 DQMLDPGPHRGDLRRARAGAAINVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPTGSV 243
 Db 181 DQMLDPGPHRGDLRRARAGAAINVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPTGSV 240
 QY 244 TELVTVTLKNSVDEINAAKASNDSPFGTDEPIVSSDIVGVSGLSDATQTKVMEVD 303
 Db 241 TELVTVTLKNSVDEINAAKASNDSPFGTDEPIVSSDIVGVSGLSDATQTKVMEVD 300
 QY 304 GSQVLKVVSVYDNEMSYTAQ 323
 Db 301 GSQVLKVVSVYDNEMSYTAQ 320
 RESULT 14
 ID Q9AJT7 PRELIMINARY; PRT; 309 AA.
 AC Q9AJT7;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
 DE (Fragment).
 GN GAP.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 CX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=clinical isolate;
 RA Amzaga M.R., Carter P.E., Cash P., McKenzie H.;
 RT "Sequence variation in mef(A) and gap genes in M-phenotype
 RT erythromycin resistant Streptococcus pneumoniae";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 DR EMBL; AJ292048; CAC27448.1; -.
 DR HSP; P00362; 1GDI.
 DR GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
 DR GO:0016491; F:oxidoreductase activity; IEA.
 DR GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR006424; GAPDH-I.
 DR InterPro: IPR000173; GAP_dhrogenase.
 DR Pfam: PF00044; gpdh; 1.
 DR Pfam: PF02800; gpdh; C; 1.
 DR InterPro: IPR00078; G3PDHGRNASE.
 DR TIGRfams; TIGR01534; GAPDH-I; 1.

DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
FT NON_TER 309 1
FT NON_TER 309 309
SQ SEQUENCE 309 AA; 32930 MW; 89E45C3BEA6A0528 CRC64;
Query Match 85.6%; Score 1468.5; DB 2; Length 309;
Best Local Similarity 91.9%; Pred. No. 1.7e-90;
Matches 285; Conservative 12; Mismatches 12; Indels 1; Gaps 1;
QY 9 GFGRIQLAFRRIONVGEVETRINDLTDPNMLAHLKYDTTQGRFDGTVEVKEGGFEVN 68
DB 1 GFGRIQLAFRRIONVGEVETRINDLTDPNMLAHLKYDTTQGRFDGTVEVKEGGFEVN 60
QY 69 GNFIKVSABRDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVVITAPGGNDV 128
DB 61 GKFIKVSABRDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVVITAPGGNDV 119
QY 129 KTVVFNTHDILDTETVIGASCTTNCLAPMAKALHDAFGI QKGLMTTTHAYTGDMIL 188
DB 120 KTVVFNTHDILDTETVIGASCTTNCLAPMAKALHDAFGI QKGLMTTTHAYTGDMIL 179
QY 189 DGPHRGDLRRARAGAAINVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPTGSGVTELVA 248
DB 180 DGPHRGDLRRARAGAAINVPNSTGAAKAIGLVIPELNGKLDGAAQRPVPTGSGVTELVA 239
QY 249 TLDKNVSVDEINAAKASNDSEFGYTEDPIVSSDIVGVSGLFDPATQTKVMEVDGSQLV 308
DB 240 VLEKNVTVDVNAAKASNESGYTEDPIVSSDIVGVSGLFDPATQTKVMEVDGSQLV 299
QY 309 KVSWSYDNEM 318
DB 300 KVSWSYDNEM 309
Search completed: May 11, 2004, 16:49:18
Job time : 52 secs

RESULT 15
Q9AJT9 PRELIMINARY; PRT; 309 AA.
AC Q9AJT9
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
DE (Fragment).
CN GAP.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Clinical isolate;
RA Amzaga M.R., Carter P.E., Cash P., McKenzie H.;
RT "Sequence variation in mef(A) and gap genes in M-phenotype
erythromycin resistant Streptococcus pneumoniae";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
DEHYDROGENASE FAMILY.
DR EMBL; AJ232046; CAC27446.1; -.
DR HSRF; P00362; IGD1.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR00173; GAP_dhdrogenase.
DR Pfam; PF00044; spdh; 1.
DR Pfam; PF02800; gpdh_C; 1.
DR PRINTS; PR00078; G3PDHGRNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
FT NON_TER 309 1
FT NON_TER 309 309

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 16:49:42 ; Search time 48 Seconds
(without alignments)
1942.967 Million cell updates/sec

Title: US-10-650-369-12

Perfect score: 1715

Sequence: 1 MVVKVINGFGRIGRLAFRR.....EMSYTAQLVRLTEYFAKIAK 336

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

```

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1715	100.0	336	9	US-09-878-766A-12
2	1715	100.0	336	10	US-09-878-781-4
3	1715	100.0	336	12	US-10-650-369-12
4	1715	100.0	336	14	US-10-134-297-4
5	1714	99.9	336	12	US-10-282-122A-74379
6	1688.5	98.5	335	10	US-09-878-781-14
7	1656.5	96.6	448	9	US-09-878-766A-22
8	1656.5	96.6	448	12	US-10-650-369-22
9	1618	94.3	336	9	US-09-878-766A-16
10	1618	94.3	336	10	US-09-878-781-8
11	1618	94.3	336	12	US-10-650-369-16
12	1618	94.3	336	14	US-10-134-297-8
13	1611	93.9	336	10	US-09-878-781-16
14	1602	93.4	336	9	US-09-878-766A-14
15	1602	93.4	336	10	US-09-878-781-6

```

16 1602 93.4 336 12 US-10-650-369-14
17 1602 93.4 336 14 US-10-134-297-6
18 1593.5 92.9 359 9 US-09-815-242-13169
19 1593.5 92.9 359 12 US-10-282-122A-74186
20 1592.5 92.9 359 9 US-09-815-242-13593
21 1584 92.4 336 9 US-09-878-766A-20
22 1584 92.4 336 10 US-09-878-781-12
23 1584 92.4 336 12 US-10-650-369-20
24 1584 92.4 336 14 US-10-134-297-12
25 1579.5 92.1 337 12 US-10-282-122A-72254
26 1578 92.0 336 9 US-09-878-766A-18
27 1578 92.0 336 10 US-09-878-781-10
28 1578 92.0 336 12 US-10-650-369-18
29 1578 92.0 336 14 US-10-134-297-10
30 1549.5 90.3 333 9 US-09-765-272-54
31 1410 82.2 336 15 US-10-369-493-18524
32 1338.5 78.0 337 15 US-10-369-493-18342
33 1337.5 78.0 333 9 US-09-815-242-10847
34 1337.5 78.0 333 12 US-10-282-122A-57294
35 1277 74.5 335 12 US-10-282-122A-53266
36 1262 73.6 336 12 US-10-282-122A-60492
37 1259.5 73.4 334 12 US-10-282-122A-51519
38 1217.5 71.0 335 12 US-10-282-122A-71734
39 1197.5 69.8 334 12 US-10-282-122A-65661
40 1196.5 69.8 357 12 US-10-282-122A-65001
41 1190 69.4 336 12 US-10-282-122A-70890
42 1183 69.0 336 9 US-09-815-242-5618
43 1183 69.0 336 9 US-09-815-242-12540
44 1183 69.0 336 12 US-10-282-122A-43997
45 993 57.9 333 12 US-10-282-122A-56828

```

ALIGNMENTS

RESULT 1

```

US-09-878-766A-12
; Sequence 12, Application US/09878766A
; Patent No. US2002004928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-09-878-766A-12

```

```

Query Match 100.0%; Score 1715; DB 9; Length 336;
Best Local Similarity 100.0%; Pred. No. 3.1e-154;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLLKYDTTQGRFDGTVEV 60
DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDPNMLAHLLKYDTTQGRFDGTVEV 60
QY 61 KEGGFVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
DB 61 KEGGFVNGNFIKVSARDPENIDWATDGVETVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTVVFNTNHDLDGTETVLSGASCTTNCLAPAKALHDAFGTKGLMTTIHA 180
DB 121 TAPGNDVKTVVFNTNHDLDGTETVLSGASCTTNCLAPAKALHDAFGTKGLMTTIHA 180
QY 181 YTGQMLDGHRRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAQRVPT 240

```

Db 181 YTGDMILDPHRRGDLRRARAGAAANIVPNSGAAKAIGLIVPELNGKLDGAAQRPVPT 240
Qy 241 GSVTELVVTLTKNSVDSDEINAAKAAASNDGFGYTEDPIVSSDIVGVSGLFDATQTKM 300
Db 241 GSVTELVVTLTKNSVDSDEINAAKAAASNDGFGYTEDPIVSSDIVGVSGLFDATQTKM 300
Qy 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 2
US-09-878-781-4
; Sequence 4, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 4
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-09-878-781-4

Query Match 100.0%; Score 1715; DB 10; Length 336;
Best Local Similarity 100.0%; Pred. No. 3.1e-154;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVVKVINGFGRIKGLAFRRIONVEGVETRIINDLTPNMLAHLKLYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIKGLAFRRIONVEGVETRIINDLTPNMLAHLKLYDTTQGRFDGTVEV 60
Qy 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
Qy 121 TAPGGNDVKTVFNTNHDILDTGTTVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 121 TAPGGNDVKTVFNTNHDILDTGTTVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Qy 181 YTGDMILDPHRRGDLRRARAGAAANIVPNSGAAKAIGLIVPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDPHRRGDLRRARAGAAANIVPNSGAAKAIGLIVPELNGKLDGAAQRPVPT 240
Qy 241 GSVTELVVTLTKNSVDSDEINAAKAAASNDGFGYTEDPIVSSDIVGVSGLFDATQTKM 300
Db 241 GSVTELVVTLTKNSVDSDEINAAKAAASNDGFGYTEDPIVSSDIVGVSGLFDATQTKM 300
Qy 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 3
US-10-650-369-12
; Sequence 12, Application US/10650369
; Publication No. US20040062774A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION

; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/10/650,369
; CURRENT FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US/09/878,766A
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 12
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-10-650-369-12

Query Match 100.0%; Score 1715; DB 12; Length 336;
Best Local Similarity 100.0%; Pred. No. 3.1e-154;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MVVKVINGFGRIKGLAFRRIONVEGVETRIINDLTPNMLAHLKLYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIKGLAFRRIONVEGVETRIINDLTPNMLAHLKLYDTTQGRFDGTVEV 60
Qy 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
Qy 121 TAPGGNDVKTVFNTNHDILDTGTTVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Db 121 TAPGGNDVKTVFNTNHDILDTGTTVISGASCTTNCLAPMAKALHDAFGIQKGLMTTIHA 180
Qy 181 YTGDMILDPHRRGDLRRARAGAAANIVPNSGAAKAIGLIVPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDPHRRGDLRRARAGAAANIVPNSGAAKAIGLIVPELNGKLDGAAQRPVPT 240
Qy 241 GSVTELVVTLTKNSVDSDEINAAKAAASNDGFGYTEDPIVSSDIVGVSGLFDATQTKM 300
Db 241 GSVTELVVTLTKNSVDSDEINAAKAAASNDGFGYTEDPIVSSDIVGVSGLFDATQTKM 300
Qy 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 4
US-10-134-297-4
; Sequence 4, Application US/10134297
; Publication No. US20030165524A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055.20
; CURRENT APPLICATION NUMBER: US/10/134,297
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 4
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-10-134-297-4

Query Match 100.0%; Score 1715; DB 14; Length 336;
Best Local Similarity 100.0%; Pred. No. 3.1e-154;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MVVKVINGFGRIKGLAFRRIONVEGVETRIINDLTPNMLAHLKLYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIKGLAFRRIONVEGVETRIINDLTPNMLAHLKLYDTTQGRFDGTVEV 60
Qy 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120

Db 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGGNDVKTWVNTNHDILDTETVTSASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180
Db 121 TAPGGNDVKTWVNTNHDILDTETVTSASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180
QY 181 YTGDMILDPGHRGGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDPGHRGGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLTKNVSVDENINAAKKAASNSFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
Db 241 GSVTELVVTLTKNVSVDENINAAKKAASNSFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
QY 301 EVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 EVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 5

US-10-282-122A-74379
; Sequence 74379, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Orlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74379
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes

Query Match 99.9%; Score 1714; DB 12; Length 336;
Best Local Similarity 99.7%; Pred. No. 3.8e-154;
Matches 335; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETRIINDLTDPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRIINDLTDPNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGGNDVKTWVNTNHDILDTETVTSASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180
Db 121 TAPGGNDVKTWVNTNHDILDTETVTSASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180
QY 181 YTGDMILDPGHRGGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDPGHRGGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTELVVTLTKNVSVDENINAAKKAASNSFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
Db 241 GSVTELVVTLTKNVSVDENINAAKKAASNSFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
QY 301 EVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 336
Db 301 EVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 336

RESULT 6

US-09-878-781-14
; Sequence 14, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SpyGapC
; OTHER INFORMATION: Protein
; US-09-878-781-14

Query Match 98.5%; Score 1688.5; DB 10; Length 335;
Best Local Similarity 98.8%; Pred. No. 1e-151;
Matches 332; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 MVVKVINGFGRIGRLAFRRIONVEGVETRIINDLTDPNMLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRIINDLTDPNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGGFEVNGNFIKVSARDPENIDWATDGVIEIVLEATGFFAKKEAAEKHLHANGAKKVI 119
QY 121 TAPGGNDVKTWVNTNHDILDTETVTSASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 180
Db 121 TAPGGNDVKTWVNTNHDILDTETVTSASCTTNCCLAPMAKALHDAFGIQKGLMTTIIHA 179
QY 181 YTGDMILDPGHRGGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
Db 181 YTGDMILDPGHRGGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 239
QY 241 GSVTELVVTLTKNVSVDENINAAKKAASNSFGYTEDPIVSSDIVGVSGLFDATQTKVM 300
Db 241 GSVTELVVTLTKNVSVDENINAAKKAASNSFGYTEDPIVSSDIVGVSGLFDATQTKVM 300

Db 240 GSVTELVTLDKNVSVDEINSAKKAASNDSPGYTEDIIVSSDIVVSGSLFDATQTKVM 299
 QY 301 EVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 336
 Db 300 EVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIAK 335

RESULT 7

US-09-878-766A-22
 ; Sequence 22, Application US/09878766A
 ; Patent No. US20020044928A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Potter, Andrew A.
 ; APPLICANT: Perez-Casal, Jose
 ; APPLICANT: Fontaine, Michael
 ; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
 ; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
 ; FILE REFERENCE: 9000-0057
 ; CURRENT APPLICATION NUMBER: US/09/878,766A
 ; CURRENT FILING DATE: 2001-09-10
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 22
 ; LENGTH: 448
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple
 ; OTHER INFORMATION: epitope fusion protein
 US-09-878-766A-22

Query Match 96.6%; Score 1656.5; DB 9; Length 448;
 Best Local Similarity 79.6%; Pred. No. 1.7e-148;
 Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;
 QY 1 MVVKVINGFGRIQLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRPDGTVEV 60
 Db 28 MVVKVINGFGRIQLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRPDGTVEV 87
 QY 61 KEGGFEVNGNFIKVSARDPE----- 81
 Db 88 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKGDFVNGKFIKVSAAE 147
 QY 82 -----NIDWATDGVIEVLE 95
 Db 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
 QY 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDGTETVISGASCTTN 155
 Db 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDGTETVISGASCTTN 267
 QY 156 CLAPWAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHEGGDLRRARAGAAANIVPNSGAA 215
 Db 268 CLAPWAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHEGGDLRRARAGAAANIVPNSGAA 327
 QY 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSELVVTLDKNVSVDEINAAKKAASNDSFGYTE 275
 Db 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSELVVTLDKNVSVDEINAAKKAASNDSFGYTE 387
 QY 276 DPVSSDIVVSGSLFDATQTKWMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 335
 Db 388 DPVSSDIVVSGSLFDATQTKWMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
 QY 336 K 336
 Db 448 K 448

RESULT 8

US-10-650-369-22
 ; Sequence 22, Application US/10650369
 ; Publication No. US20040062774A1
 ; GENERAL INFORMATION:

; APPLICANT: Potter, Andrew A.
 ; APPLICANT: Perez-Casal, Jose
 ; APPLICANT: Fontaine, Michael
 ; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
 ; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
 ; FILE REFERENCE: 9000-0057
 ; CURRENT APPLICATION NUMBER: US/10/650,369
 ; CURRENT FILING DATE: 2003-08-27
 ; PRIOR APPLICATION NUMBER: US/09/878,766A
 ; PRIOR FILING DATE: 2001-09-10
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 22
 ; LENGTH: 448
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple
 ; OTHER INFORMATION: epitope fusion protein
 US-10-650-369-22

Query Match 96.6%; Score 1656.5; DB 12; Length 448;
 Best Local Similarity 79.6%; Pred. No. 1.7e-148;
 Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;
 QY 1 MVVKVINGFGRIQLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRPDGTVEV 60
 Db 28 MVVKVINGFGRIQLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRPDGTVEV 87
 QY 61 KEGGFEVNGNFIKVSARDPE----- 81
 Db 88 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKGDFVNGKFIKVSAAE 147
 QY 82 -----NIDWATDGVIEVLE 95
 Db 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
 QY 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDGTETVISGASCTTN 155
 Db 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDGTETVISGASCTTN 267
 QY 156 CLAPWAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHEGGDLRRARAGAAANIVPNSGAA 215
 Db 268 CLAPWAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHEGGDLRRARAGAAANIVPNSGAA 327
 QY 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSELVVTLDKNVSVDEINAAKKAASNDSFGYTE 275
 Db 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSELVVTLDKNVSVDEINAAKKAASNDSFGYTE 387
 QY 276 DPVSSDIVVSGSLFDATQTKWMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 335
 Db 388 DPVSSDIVVSGSLFDATQTKWMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
 QY 336 K 336
 Db 448 K 448

RESULT 9

US-09-878-766A-16
 ; Sequence 16, Application US/09878766A
 ; Patent No. US20020044928A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Potter, Andrew A.
 ; APPLICANT: Perez-Casal, Jose
 ; APPLICANT: Fontaine, Michael
 ; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
 ; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
 ; FILE REFERENCE: 9000-0057
 ; CURRENT APPLICATION NUMBER: US/09/878,766A
 ; CURRENT FILING DATE: 2001-09-10
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 16
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-09-878-766A-16

Query Match
  94.3%; Score 1618; DB 9; Length 336;
Best Local Similarity 92.9%; Pred. No. 5.1e-145;
Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MVVKVINGFGRIGRLAFRIQNVGEVETRIINDLTPNNLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRIQNVGEVETRIINDLTPNNLAHLKYDTTQGRFDGTVEV 60

Qy 61 KEGFEVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKAAEKLHANGAKKVI 120
Db 61 KEGFEVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKAAEKLHANGAKKVI 120

Qy 121 TAPGNDVKTWVFNTHDILDGTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 180
Db 121 TAPGNDVKTWVFNTHDILDGTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 180

Qy 181 YTGDMILDGPGRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAARVPVPT 240
Db 181 YTGDMILDGPGRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAARVPVPT 240

Qy 241 GSVTELVVTLDKNVSVDENAMKAASNDSPGYTEDPIVSSDIVGSGSLFDATQTKVM 300
Db 241 GSVTELVVTLDKNVSVDENAMKAASNDSPGYTEDPIVSSDIVGSGSLFDATQTKVM 300

Qy 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336

Qy 301 TVDGNQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336
Db 301 TVDGNQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336

RESULT 10
US-09-878-781-8
; Sequence 8, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
; FILE REFERENCE: 9000-0035
; CURRENT APPLICATION NUMBER: US/09/878, 781
; PRIOR FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 8
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-09-878-781-8

Query Match
  94.3%; Score 1618; DB 10; Length 336;
Best Local Similarity 92.9%; Pred. No. 5.1e-145;
Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MVVKVINGFGRIGRLAFRIQNVGEVETRIINDLTPNNLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRIQNVGEVETRIINDLTPNNLAHLKYDTTQGRFDGTVEV 60

Qy 61 KEGFEVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKAAEKLHANGAKKVI 120
Db 61 KEGFEVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKAAEKLHANGAKKVI 120

Qy 121 TAPGNDVKTWVFNTHDILDGTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 180
Db 121 TAPGNDVKTWVFNTHDILDGTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 180
```

```
Qy 181 YTGDMILDGPGRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAARVPVPT 240
Db 181 YTGDMILDGPGRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAARVPVPT 240

Qy 241 GSVTELVVTLDKNVSVDENAMKAASNDSPGYTEDPIVSSDIVGSGSLFDATQTKVM 300
Db 241 GSVTELVVTLDKNVSVDENAMKAASNDSPGYTEDPIVSSDIVGSGSLFDATQTKVM 300

Qy 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336

Qy 301 TVDGNQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336
Db 301 TVDGNQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336

RESULT 11
US-10-650-369-16
; Sequence 16, Application US/10650369
; Publication No. US20040062774A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/10/650,369
; PRIOR FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US/09/878, 766A
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 16
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-10-650-369-16

Query Match
  94.3%; Score 1618; DB 12; Length 336;
Best Local Similarity 92.9%; Pred. No. 5.1e-145;
Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MVVKVINGFGRIGRLAFRIQNVGEVETRIINDLTPNNLAHLKYDTTQGRFDGTVEV 60
Db 1 MVVKVINGFGRIGRLAFRIQNVGEVETRIINDLTPNNLAHLKYDTTQGRFDGTVEV 60

Qy 61 KEGFEVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKAAEKLHANGAKKVI 120
Db 61 KEGFEVNGNFIKVSARDPENIDWATDGEIVLEATGFFAKKAAEKLHANGAKKVI 120

Qy 121 TAPGNDVKTWVFNTHDILDGTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 180
Db 121 TAPGNDVKTWVFNTHDILDGTETVSGASCTTNCCLAPMAKALHDAFGIQKGLMTTIIA 180

Qy 181 YTGDMILDGPGRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAARVPVPT 240
Db 181 YTGDMILDGPGRGDLRRARAGANIVPNSTGAAGAIGLVIPELNGKLDGAARVPVPT 240

Qy 241 GSVTELVVTLDKNVSVDENAMKAASNDSPGYTEDPIVSSDIVGSGSLFDATQTKVM 300
Db 241 GSVTELVVTLDKNVSVDENAMKAASNDSPGYTEDPIVSSDIVGSGSLFDATQTKVM 300

Qy 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336
Db 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336

Qy 301 TVDGNQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336
Db 301 TVDGNQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336

RESULT 12
US-10-134-297-8
; Sequence 8, Application US/10134297
; Publication No. US20030165524A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
```


APPLICANT: Potter, Andrew A.
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
FILE REFERENCE: 9000-0055.20
CURRENT APPLICATION NUMBER: US/10/134,297
CURRENT FILING DATE: 2002-04-26
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 8
LENGTH: 336
TYPE: PRT
ORGANISM: Streptococcus uberis
US-10-134-297-8

Query Match 94.3%; Score 1618; DB 14; Length 336;
Best Local Similarity 92.9%; Pred. No. 5.1e-145;
Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

QY 1 MVVKGINGFGRIGRLAFRRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
DB 1 MVVKGINGFGRIGRLAFRRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
DB 61 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVFNTNHDILDTETVSGASCTTNCNCLAPMAKALHDAFGIOKGLMTTIIA 180
DB 121 TAPGNDVKTIVFNTNHDILDTETVSGASCTTNCNCLAPMAKALHDAFGIOKGLMTTIIA 180
QY 181 YTGDMILDGPHRGDLRRARAGAAANIPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
DB 181 YTGDMILDGPHRGDLRRARAGAAANIPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTLELVTLTKNVSVDENINAAKAAANDSFYTEDPIVSSDIVGVSGLFDATQTKVM 300
DB 241 GSVTLELVTLTKNVSVDENINAAKAAANDSFYTEDPIVSSDIVGVSGLFDATQTKVM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336
DB 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336

RESULT 13
US-09-878-781-16
Sequence 16, Application US/09878781
Publication No. US20030082781A1
GENERAL INFORMATION:
APPLICANT: Bolton, Alexandra J.
APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
APPLICANT: Potter, Andrew A.
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
FILE REFERENCE: 9000-0055
CURRENT APPLICATION NUMBER: US/09/878,781
CURRENT FILING DATE: 2002-09-10
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 16
LENGTH: 336
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SeqGapC
US-09-878-781-16

Query Match 95.9%; Score 1611; DB 10; Length 336;
Best Local Similarity 95.2%; Pred. No. 2.3e-144;
Matches 320; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

QY 1 MVVKGINGFGRIGRLAFRRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60

DB 1 MVVKGINGFGRIGRLAFRRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
DB 61 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVFNTNHDILDTETVSGASCTTNCNCLAPMAKALHDAFGIOKGLMTTIIA 180
DB 121 TAPGNDVKTIVFNTNHDILDTETVSGASCTTNCNCLAPMAKALHDAFGIOKGLMTTIIA 180
QY 181 YTGDMILDGPHRGDLRRARAGAAANIPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
DB 181 YTGDMILDGPHRGDLRRARAGAAANIPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTLELVTLTKNVSVDENINAAKAAANDSFYTEDPIVSSDIVGVSGLFDATQTKVM 300
DB 241 GSVTLELVTLTKNVSVDENINAAKAAANDSFYTEDPIVSSDIVGVSGLFDATQTKVM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336
DB 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336

RESULT 14
US-09-878-766A-14
Sequence 14, Application US/09878766A
Patent No. US20020044928A1
GENERAL INFORMATION:
APPLICANT: Potter, Andrew A.
APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
FILE REFERENCE: 9000-0057
CURRENT APPLICATION NUMBER: US/09/878,766A
CURRENT FILING DATE: 2001-09-10
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 14
LENGTH: 336
TYPE: PRT
ORGANISM: Streptococcus agalactiae
US-09-878-766A-14

Query Match 93.4%; Score 1602; DB 9; Length 336;
Best Local Similarity 91.7%; Pred. No. 1.7e-143;
Matches 308; Conservative 17; Mismatches 11; Indels 0; Gaps 0;

QY 1 MVVKGINGFGRIGRLAFRRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
DB 1 MVVKGINGFGRIGRLAFRRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY 61 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
DB 61 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEATGFFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVFNTNHDILDTETVSGASCTTNCNCLAPMAKALHDAFGIOKGLMTTIIA 180
DB 121 TAPGNDVKTIVFNTNHDILDTETVSGASCTTNCNCLAPMAKALHDAFGIOKGLMTTIIA 180
QY 181 YTGDMILDGPHRGDLRRARAGAAANIPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
DB 181 YTGDMILDGPHRGDLRRARAGAAANIPNSTGAAGAIGLVIPELNGKLDGAAQRPVPT 240
QY 241 GSVTLELVTLTKNVSVDENINAAKAAANDSFYTEDPIVSSDIVGVSGLFDATQTKVM 300
DB 241 GSVTLELVTLTKNVSVDENINAAKAAANDSFYTEDPIVSSDIVGVSGLFDATQTKVM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336
DB 301 EVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIAK 336

RESULT 15
US-09-878-781-6
; Sequence 6, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-09-878-781-6

Query Match 93.4%; Score 1602; DB 10; Length 336;
Best Local Similarity 91.7%; Pred. No. 1.7e-143;
Matches 308; Conservative 17; Mismatches 11; Indels 0; Gaps 0;

Qy 1 M V V K V G I N G E R I G R L A P R I Q N V E G V E V T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 60
Db 1 M V V K V G I N G E R I G R L A P R I Q N V E G V E V T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 60

Qy 61 K E G G F E V N G F I K V S A E R D P E N I D W A T D G V E I V L E A T G F F A K E A E K H L H A N G A K K V I 120
Db 61 K E G G F E V N G Q F V K V S A E R E P A N I D W A T D G V E I V L E A T G F F A S K E K A G O H I H E N G A K K V I 120

Qy 121 T A P G N D V K T V V F N T N H D I L D G T E T V I S G A S C T T N C L A P N A K A L H D A F G I Q K G L M T T I H A 180
Db 121 T A P G N D V K T V V F N T N H D I L D G T E T V I S G A S C T T N C L A P N A K A L Q D N F G V K Q G L M T T I H A 180

Qy 181 Y T G D M I L D G P H R G G D L R R A G A A N I V P N S T G A A K A I G L V I P E L N G K L D G A A Q R V P V T 240
Db 181 Y T G D M I L D G P H R G G D L R R A G A A N I V P N S T G A A K A I G L V I P E L N G K L D G A A Q R V P V T 240

Qy 241 G S V T E L V V T L D K N V S D E I N A A K A A S N D S F G Y T E D P I V S S D I V G V S Y G S L F D A T Q T K Y M 300
Db 241 G S V T E L V A T L E K D V T V E V N A A K A A N D S Y G Y T E D P I V S S D I V G I S Y G S L F D A T Q T K Y Q 300

Qy 301 E V D G S Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A K 336
Db 301 T V D G N Q L V K V S W Y D N E M S Y T S Q L V R T L E Y F A K I A K 336

Search completed: May 11, 2004, 16:55:20
Job time : 49 secs